



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Social Networks 27 (2005) 139–153

SOCIAL
NETWORKS

www.elsevier.com/locate/socnet

Who benefits from network analysis: ethics of social network research[☆]

Charles Kadushin^{*}

*Cohen Center for Modern Jewish Studies and Heller School of Social Policy and Management,
Brandeis University, Waltham, MA 02154, USA*

Abstract

The success of social network research (SNR) has led to expectations that in addition to academic research, SNR can introduce people to one another, solve organizational problems, map the epidemiology of AIDS, and catch criminals and terrorists. Since SNR requires that names of both respondents and their contacts be collected and used in most analyses, Institutional Review Boards become very concerned. Experiences of the author, participants in the 2003 Sun Belt Conference and the Social Network List Serve illustrate ethical issues. Proper handling of the data and the analysis, including complete control by the investigator can virtually eliminate harm to respondents and those they nominate, though perhaps not to the satisfaction of IRBs. On the benefit side, academic researchers always benefit, organizations, society and science may benefit, but individual respondents rarely do. © 2005 Elsevier B.V. All rights reserved.

Keywords: Social network research; Social research methods; Ethics; Institutional review boards

1. Introduction

The social network field may have become a victim of its own successes. The mapping of social networks with names, dates and places has become a major industry. Barry Wellman reports that “Business 2.0 anointed ‘social network applications’ in 2003 as ‘The Technology

[☆] Appreciation for comments and ideas to the Sunbelt 2003 session on network ethics, to Ron Breiger, Scott Feld, Joe Labianca, and the Social Networks List Serve. This paper is a personal statement, however, and I take full responsibility for its content and positions on ethics.

^{*} Tel.: +1 781 736 2060; fax: +1 781 736 3929.

E-mail address: kadushin@brandeis.edu.

of the Year” (Wellman, 2003). The New York Times has celebrated social networks as one of the “new ideas” of the year (Gertner, 2003). PC Magazine reviewed five Internet sites that attempt to introduce people to one another.

Introduced by Stanley Milgram in 1967, the theory of six degrees of separation, which supposes that you’re just a half-dozen introductions away from anyone you want to meet, has found the Internet. Sites like LinkedIn, which take hold of the six-degrees concept and put it to practical use, let you take advantage of chains of acquaintances to contact people down the line. They’re known as social-networking services. Such services use the Internet to help users expand their networks of personal and business relationships.

The process is simple. After joining one or more of these sites, you send messages to people you know, asking them to join. They in turn invite people they know, and so on. In this way, you construct an enormous network of people to whom you have personal links.

... Friendster boasts three million users; none of the others has even approached 100,000 (Metz, 2004), p 131.

One of these sites has applied for a patent. The value of one’s network of friends can be calculated.¹ At the other extreme there are maps of terrorist networks and firms that apparently make a living by providing crime fighting units with software to map criminal networks. Saddam Hussein was said to have been captured in part through the application of social network mapping (Fassihi, 2003). Many organizations attempt to improve their efficiency through sociometric analyses (Krebs, 2003). Research and development laboratories map major gatekeepers of critical information. There are maps of who works with whom in biotechnology. Epidemiology was founded on the tracing of agents who carried disease and modern network methods have been applied to the HIV positive field. Structures of national leaders and decision-makers have been studied, as well as the structure and function of corporate overlaps. One could go on and on and produce what is essentially a bibliography of important social network studies.

The ethical issues are both straightforward and complex. In standard practice social science research, anonymity and confidentiality are both routinely granted to respondents, informants, and subjects in experiments and observations. In large-scale survey research with at least several hundred respondents these guarantees are very easy to keep. The researcher has no interest in the particular names of respondents, except in the case of panel studies when prior respondents need to be contacted again. Looking them up serves no purpose whatsoever.² In smaller scale qualitative studies, often organization or small community studies, who are the respondents even when given promises of anonymity may be

¹ The authors are apparently unaware of Burt’s work (Teten et al., 2004).

² In the days before computer databases, it was almost totally impractical to check up on any one respondent. Now, with computers and linked files, it is easy enough if the links can be made, but from a practical point of view still useless to the researcher, though valuable to a law enforcement agency that may wish to discover who, for example, are using illegal drugs, if that is the topic of the survey.

obvious to both potential readers and to the social scientists. The latter often cannot successfully analyze the data without knowing who the respondents are. Eventual publication usually involves changing the names of respondents as well as information such as their age and occupation that might give them away even though some of this background of subjects or respondents may be important to the narrative. When it comes to the comparative organization or community studies there are further difficulties because it may be impossible or even undesirable for analytic purposes to disguise the names of the organizations or communities. Organization consultants, whether academically based or not, further face the dilemma that the observations or surveys that they produce may have consequences for the individuals surveyed or observed that the subjects may not have been aware of (Borgatti and Molina, 2003).

An extreme, but clever new technique in organizational network studies utilizes real time data. An article in *Technology Review* offers the following by an innovative network researcher (Pentland, 2004):

“Who are the experts within your organization? Who has the most decision-making influence? Recently, managers have started mining data from e-mail, Web pages, and other digital media for clues that will help answer such questions. That’s a start, says MIT Media Lab researcher Alex Pentland, but it misses the real action: studies of office interactions indicate that as much as 80 percent of work time is spent in spoken conversation, and that critical pieces of information are transmitted by word of mouth in a serendipitous fashion. Fortunately, the data infrastructure for mining real-world interactions is already in place. Most working professionals already carry microphones (cell phones), and many also carry PDAs with ample computational horsepower. This foundation of mobile communications and processing power will support an exciting new suite of business applications: reality mining.”

In addition to these issues that are well known but for which there may not necessarily be obvious solutions, social network data have one troublesome and distinctive attribute: the collection of names of either individuals or social units is not incidental to the research but its very point. Further, the network analyst in collecting information about who relates to whom is not confined to the names of respondents or informants within the study, for they may give the names of others who have no idea that they are being named (Borgatti and Molina, 2003). In studies of elites, for example, it is common practice for a connection to be considered as valid between two individuals not within the original list of respondents if they are both named by a respondent within the study (Alba and Kadushin, 1976; Alba and Moore, 1978; Higley and Moore, 1981; Moore, 1979; Moore, 1961). The relationship or tie between individuals may be obviously pertinent to the topic of the investigation, say influence on public policy, but may also be less obviously related and even more “private” such as who is friends with whom (Kadushin, 1995). In organization settings, while who works with whom might seem a legitimate topic, who is friends with whom may lie outside the purview of an employer (Borgatti and Molina, 2003).

Finally, one need not go to respondents and ask them about their relationships. Important work in social network analysis has been done with public databases, or archived material that could be turned into databases. For example, Valdis Krebs has published a map of relations between Internet companies based on public data

<http://www.orgnet.com/netindustry.html>. More controversially, he has analyzed networks of 9/11 terrorists based on publicly available data (Krebs).³ Recently there has been concern about using network techniques in interrogations in Iraq.

There have been a number of studies of corporate overlap based on publicly available records, for example (Mintz and Schwartz, 1985).⁴ Network analysis always makes visible that which cannot be seen by the naked eye. In the case of analysis using public records, the data were already there for “all to see” but in fact, without first collecting data from various sources and putting them in a data base and only then analyzing and graphing them, the data would have remained invisible. If these data are public, at least in principle, is there an ethical issue? The ethics lie in making assertions about connections, centrality, and power when there may be some question about the reliability and accuracy of the sources of the data even when they are publicly available. Perhaps asserting that IBM has a central position in the network of computer firms in the Internet is one thing, but asserting that a given named person has a central role in a terrorist network (and thereby marking the person for arrest or even “elimination”) is another. Granting that this is an extreme example, though one that is now on the agenda, what about making visible the connection of a particular individual that may have negative (or positive) consequences for that individual? The person in question may never have expected that the available data be used in this way.

Epidemiology in some ways can be said to have “invented” network analysis in the 19th Century. Currently, it has wide use in tracing illnesses and more recently has been used in AIDS research, subject to all sorts of safeguards that Klovdahl discusses (Klovdahl, 2005). The fact that HIV status is often hidden for good reasons raises a number of problems (Shelley et al., 1995). Contact tracing had been the *modus vivendi* of work with TB and even venereal diseases. In the beginning, HIV status was tantamount to sentencing an individual to a short and nasty life and resulted in much stigma. As a result, in the early years of the discovery of the disease it was virtually impossible to introduce studies that rested on contact tracing. GMHC [Gay Men’s Health Crisis], for example, was strongly opposed to such work. The organization currently notes in advice to on its Web site that “New York State’s HIV Confidentiality law prohibits health care providers and social service providers from disclosing your HIV status without your written consent.”

There is another point of view, rooted in the history of the social network field. The inventor of “sociometry,” Jacob Moreno, insisted that sociometric data were not valid *unless* the subjects knew that their answers would have consequences. That was the principle behind his and Jennings famous study of the cottage system of delinquent women teenagers. One of the first applications of network analysis was in the early 30’s when “sociometry” showed how artificially to construct primary groups of adolescent young women who were incarcerated in an institution that housed them in separate cottages. The aim was to place women who liked one another into the same cottage rather than having cottages populated by

³ But Krebs in an email to the social network list (13 July 2004) has observed: “Of the many social network analysts that have looked at finding the ‘terrorist needle’ in the ‘population haystack’ not many are convinced that it can be done. Most are convinced that you will first drown in false positives. There are exceptions . . . but they may have a ‘silver bullet’ for sale. There are other methods than examining a billion straws. . . . “If it was as easy as. . . (1) ask 10 questions (2) mine a billion records (3) spit out 19 hijackers it would have been done.”

⁴ For a review, see (Mizruchi, 1996).

warring cliques (Moreno, 1953). When grouped into cottages with greater group cohesion, the women were less likely to be cantankerous. Moreno successfully constructed primary groups or natural cliques or what were described above as “strong ties” that substituted for the lack of functioning families in this situation. Since Moreno, the general tendency in the network field has been not to reveal individual identities. The famous “Bank Wiring Room” omitted details that would have revealed the identity of the subjects (Roethlisberger and Dickson, 1939). An immediate practical issue is the reaction of literate “natives” to the revelation of the details of their interrelations.

For academics, the abstract ethical issues are overshadowed by the bureaucratic need to submit all research before it proceeds to a Protection of Human Subjects Institutional Review Board (IRB).⁵

The fundamental issue is who benefits from the network analysis? This is another version of the cost benefit analysis so dear to the hearts of Human Subject Review Committees who often conclude in the case of network research that the hazards to the individual outweigh any benefits to “Science.” Further, many IRB’s conclude that *any* way of identifying individual participants is illegitimate and unethical. When viewed in this way, with privacy as an absolute right, network data collection is harmful regardless of any possible benefits. When “Science” is not involved and IRB’s are not consulted, as in the examples of interrogation of suspected terrorists, or in the discovery within an organization who are the “real” leaders, or when individuals stand to gain dates, business contacts or sales advantages from the linking of names, the issues become even more murky. The concern of “legitimate” network researchers is that concern over individual rights to privacy will ultimately torpedo the entire social network field. On the other hand, an IRB, as will be seen, can provide guarantees for respondents and an institutional structure that enforces these guarantees. This situation may actually aid in the willingness of respondents to participate. What follows is a more detailed discussion of these issues, based in part, by the contributions made by network researchers at the Sunbelt, 2003 and by various comments contributed to the Social Network List Serve.

2. Personal experiences

I begin with a few personal examples. I follow this method because my experiences were mirrored by many in the in the 2003 Sunbelt session on ethics, but the details of these examples are more familiar to me. Following, I will discuss other issues that emerged from that discussion or various postings to the Social Networks List Serve. I will not cover surveys of ego networks in which names are not preserved but rather are used only to assist respondents in talking about members of their first order zone.

2.1. School surveys

One of my first experiences was in a study of Juilliard School of Music in the early 60s (Kadushin, 1969). The intent was not to study ego networks but to use the data to construct a network of the entire school, though we would have been satisfied with the characteristics

⁵ Ironically, the development of such boards as a Federal requirement was stimulated in part by work of Stanley Milgram—not his small world study, but his famous study of authoritarianism.

of the ego network that would have been obtained by linking questionnaire responses to the named persons. We asked, in mailed questionnaire, who are your “Three best friends” [the format used in those days]. This question apparently was responsible for a rumor that I was trying to discover who was homosexual. Response rates in some departments suffered and the question was unanalyzable because of extensive missing data. This was before the days of IRBs. In retrospect, the IRB apparatus, had they approved of the method, might have alleviated some suspicions, but perhaps not. It is one thing to gather this sort of data from school children, but another from young adults.

2.2. *Research on elites*

In the late 60’s I worked with a team from Columbia’s Bureau of Applied Social Research and the Yugoslav Opinion Research Center on studies of elites in Yugoslavia. Later we were involved together with colleagues in Czechoslovakia in studying the Czech elite. In the same time period I studied the American intellectual elite (Kadushin, 1974). They all contained data on who the respondent talked with about what. These data were used to develop very large computer constructed sociograms (Alba and Guttman, 1972; Alba, 1973). Of course we needed the names to construct the sociograms. During the ’68 student “revolution” on the Columbia Campus, I took all the data home for safekeeping, since I was fearful that the Yugoslavia data together with the names would be publicly circulated. A book was eventually published, with analyses of the characteristics of different kinds of Yugoslav elite circles, but the sociograms were not (Kadushin and Abrams, 1973a,b). When the Russians reentered Czechoslovakia in August after the famous democratic spring of 1968, the data were temporarily placed in a safe deposit box.⁶ Later we collected similar data on the general American elite as well as the Australian elite and the German elite (Higley et al., 1991). Oddly enough, elected officials are in a special “exempt” category—subject to review by IRBs but research involving them is routinely waived from the usual restrictions on confidentiality. Of course, many elites are not elected. My most recent foray into elite studies was collecting data on the French financial elite in 1990 (Kadushin, 1995).

In none of these did we have to contend with an IRB. But we did have potentially sensitive if not dangerous data. Of course people could dissemble about their contacts: the American elite data was conducted during Watergate, though the matter had not yet broken. One of the key figures in the scandal was interviewed: he said that his major concerns were with issues pertaining to the environment and he gave us a list of contacts in that field. Needless to say, his contacts involved in trying to cover up the scandal were not mentioned and we did not even know at the time that it was on the agenda. Sometimes, things are just too “hot” even to attempt a network study. I don’t know if we would have interviewed some of the people we did had we known about Watergate at the time. I did withdraw my proposal to study the American intellectual elite when I first proposed it in the mid sixties because there was a Senate investigation into the loyalty of prominent American intellectuals who were beginning to be opposed to the war and I wanted to use the war as a network “tracer.”

The investigation was dropped and I submitted the proposal a year later. It was eventually funded by NSF and became *The American Intellectual Elite* (Kadushin, 1974). In the book I

⁶ The box was subsequently retrieved and the raw data, on punch cards, sits in my home office, as yet unanalyzed.

published a list that summed up the top 70 names mentioned on seven network questions that asked who influenced the respondent and who were generally most influential on various topics. The names were listed in four ranks, the top 10, the next 10, ranks 21–25 and 26 through 27. The latter two sets had numerous ties. Within each set of ranks names were listed alphabetically so as to minimize invidious comparisons. Among the top ten was Dan Bell, who was then noted by careless reviewers as the leading American intellectual, though among the top ten he was leading only by virtue of the alphabet. Since interviews and links between people were anonymous, the intent of the list was to indicate the kind of people being studied. I do not think the list violated any privacy concerns, but it retrospect it was at the least an intellectual error. Readers and reviewers paid more attention to the list than to the contents of the book. I would not (and did not) repeat this mistake. A sociometric list can be characterized by presenting the attributes that correlate with leadership rather than by presenting the names. This is a good rule, both from an intellectual and ethical standpoint. But it was not completely followed.

Having the names on a sociogram, however, helps the analyst interpret the data, despite the social network field's interest in formal propositions based on pure structure. Which people are in which positions can give the analyst who knows the original scene many clues as to the formal properties of the network or the attributes of the clusters that can then be formally tested. The names are then removed for publication. In the computer generated sociogram on intellectual circles as published in the *American Intellectual Elite*, (Fig. 1, p. 85), the people (or their intellectual or biological children) clustered on the right of sociogram years later turned out to be the core of the neo-conservatives. Their placement on the right of sociogram was of course pure accident. Since I am the only one who knows who they were, I cannot publicly make such a claim since their attributes alone do not entirely predict their politics 25 years subsequent to the sociogram. But their personal associations as depicted in the network turned out to be excellent predictors. Even now, however, the sociogram as printed in the book is not clear enough for anyone to track down who was related to whom. My private copy is much blown up and clearer.

My study of the French financial elite posed several ethical problems. The research was done at the request of a journalist who had access to elite French circles. We made a deal: she would conduct the interviews according to my specifications with respect to the network questions and the sample design; I would analyze the data for her in return for the right to publish the data in an academic journal while she would publish the data as a book. There was no IRB review. Her access meant that we could ask whom the respondent socialized with—critical for this population. We agreed that interviewees would be promised anonymity at least for the network part of the interview. But when it came time to publish the book, the journalist wanted to put names to the nodes in the sociogram. We had quite a row. The list of names was important to her, since the book was intended as a description of the French financial institutions and the way they worked. Who made what kinds of deals with whom, most of which were part of the record in any case, constituted an important dimension of the book and the publisher would not accept it otherwise. We compromised: the names as located on the sociogram were printed in the book, but none of the edges were, and thus who was connected with whom could not be determined, though the general clustering of persons could be adduced. It was assumed that the readers of the book, bankers and investment capitalists in the Anglo Saxon world, would never read the *American Sociological Review*

(Kadushin, 1995). The general rule for network diagrams is that the labels are often a numerical ID starting with 1 and ending with N, as recommended by Steven Borgatti (in exchanges on the Social Networks List) and as automatically implemented in his UCINET (Borgatti et al., 2004). Even numerical ID's can be problematic in that for most research a key must link the number to the actual name so that data can be tracked. There is an elaborate protocol required by the Federal government for handling sensitive data that involves several linked keys accessible only in an encrypted, locked and protected computer in a manner especially reviewed and approved. In the published article I did not use numerical ID's but instead added information for each node about their political party orientation (four possibilities) and whether they were graduates of the most relevant grand école. Any number of people could fit the combination. When Frank (Frank and Yasumoto, 1998) asked for the data set so that he could reanalyze it and add many details about corporate relations and deal making that he then obtained through laborious research of his own, it was apparent that he could easily link the names to the sociogram. We agreed that he would not do so, even though this made for some awkwardness in the analysis.

2.3. *Organization research*

Organization network research can be especially vexing from an ethical standpoint because people's jobs and futures may be directly at stake. Less problematic is research conducted entirely for academic purposes. Oddly, purely academic network organizational research is relatively rare, since the access required to the organization usually requires a quid pro quo in that the organization usually requires a report which may aid them in understanding and possibly reorganizing their organization. *Management and the Worker* (Roethlisberger and Dickson, 1939) set the standard by not revealing any names, but they were dealing with lower level factory employees who in some way could be regarded as interchangeable. Organization research that I have personally conducted has always been as a consultant; this status has granted me access, under varying conditions. But there are conditions upon which I will not compromise: the data are always under my direct control, must be collected under guidelines that I describe, must reside on my computers, as do the names associated with the data. Confidentiality is always guaranteed. The data are never the property of the firm for whom I am a consultant. Names are never associated with network graphs or with network indices and are never revealed to either management or employees. Rather, general patterns are described and used to suggest they way things currently flow and how matters might be changed. "Things" as flows depend on the purpose of the investigation—communication, prestige, authority, and even friendship. If the organization cannot meet these conditions, then they must look elsewhere for someone to carry out their investigation.

Typically, these conditions cannot be met by classified or military research and so I do not do this kind of work. In the early 1970's, a research firm that had engaged in classified research for the intelligence community wished to expand its work into the social network field and approached me through a mutual friend. With the aid of NSF and Department of Education grants we had developed several social network programs capable of handling very large matrixes (Alba and Guttman, 1972; Alba, 1973) and I suppose this made me attractive. I turned them down. In the late 70's I inquired via a colleague well connected

to the intelligence community into progress the intelligence services might have made in handling large social networks. I had hoped that they just might share some of their technical computer work or at least some general ideas. Our own work was of course in the public domain. My colleague reported that for the first time in his life, he met a stonewall. I suppose the intelligence community had indeed made some progress but did not want anyone to know about it. This one-way street is not very useful for scientific research nor, ultimately, for progress and so I believe, unethical, though others equally concerned with our national safety may disagree with me.

There are some borderline issues in the organizational research typical of the genre. Most vexing in any organizational research whether it be the standard employee surveys or network studies is the size of the collectivity identified, whether its name is used, and whether statuses are named. Employee surveys are typically used to identify morale problems, and serve as two-way communication: employee problems and concerns are sent upward to management, and management discovers whether its messages and programs are being heard and absorbed by employees. In a “best practices” use of employee surveys, work-units are assembled and fed back the results of the survey that are then discussed. Members of the work unit can see where they stand vis-à-vis other units of the company, whether or not there is consensus on their problems and views, and if there is relative consensus whether airing these views alleviates a problem of “pluralistic ignorance” (Merton, 1968)⁷ in which persons think that a view that is actually widely shared is held only by them. The unit can then take steps to try to solve problems that have been identified. A belief that the survey will result in action improves response rates the next time the survey is administered. On the other hand, an obvious problem in this “best practice” is the size of the unit identified: if it is too small, then the survey ceases to be anonymous. If it is too large, then the practical value of the feedback is reduced. Typically organizations have a policy about the size of the reporting unit; it is often no smaller than 25 persons. But naming the unit can point fingers at those in charge of the unit and the data can be used to sanction them positively or negatively. I have strongly objected to the use of survey data as a component in a salary or appointment reward system. Generally, I have been successful since it is easy to point out that tying survey results to personnel actions can result in less than truthful responses, thus obviating the entire exercise. There is a similar problem with identifying statuses, especially in network research. There may be only a few secretaries or vice-presidents, etc., in a particular network and showing results for them may identify them. This is a typically difficult dilemma and one reason why I like to be in total control of the data as well as the feedback of social network data.

Another dilemma is to whom to give the feedback and at what level of detail. I prefer to make reports that are posted so that all who participated in the study, as well as management, can see them. Such a posting may limit what can be said. A borderline issue is whether to make more detailed reports to one’s fellow consultants in a team consulting effort so that we may use that information in a general assessment of an intervention. I tend to be much more free in giving information and discussing results with my consulting team—who are after all bound by the same confidentiality agreements as I am—than with management. When

⁷ The original publication date was 1957. This concept has been widely cited and attributed to various authors, but Merton was the first to coin the name.

the team eventually reports to management, “incriminating” details are removed or put in general terms. If a consultant thinks that a particular person is a problem, that can be shared and usually with the person themselves, but in a sensitive way that does not compromise other people’s confidentiality and rights. These feedback issues intersect with the named status identification dilemma. In one network study of an industrial setting, enlisting the union was the only way to get the high response rate required by network studies. And in general, if the organization being studied has a union, it is imperative to get their cooperation in a network study. In this case the union cooperated because they were concerned about the effectiveness of their shop stewards. We promised a general report about the centrality or lack of it of the union shop stewards in various flows. As it turned out from the union’s point of view, the news was good: the shop stewards were far more central than the union had imagined.

3. Further issues

There are other important issues that were brought up in the Sunbelt Workshop. One is the matter of the accuracy of information gathered by others, sometimes for public use but often for surveillance purposes; another was the issue of second parties—persons or organizations who were not respondents to a network investigation but whose names came up in the course of the investigation and who may have been used to link names of respondents; finally, the entire matter of IRBs who were most often seen as a serious obstacle. I will deal with each issue in turn.

3.1. Accuracy of data

Problems in this area were one of the motivations for our concern with social networks ethics in the first place. The problem is especially vexing when the analyst has not collected the data and thus has little idea of its validity. When the data are used to analyze cut-points in a network such that eliminating a node may break the network apart, inaccuracies can literally be fatal to an “innocent” person. This is not the forum to debate the ethics of counter-terror activities, but few are comfortable with non-judicial killing of people who are not combatants or key terrorists. The issues were well expressed by Elin Waring in a Social Network List Serve posting on August 5th, 2004. Professor Waring is an expert on criminal social networks and has published articles and books in this area (Finckenauer and Waring, 1997; Finckenauer and Waring, 1998; Waring, 2002). She observes, “I personally think that networks are the way to look at the organization of crime, and many SNA tools are helpful for this, but I would be very, very cautious.”

She further notes,

“The issue with most organized crime and terrorist network analysis, as has already been pointed out, is that *the data that are available tend to reflect the story that particular actors, usually prosecutors or reporters, want to tell* [italics added]. So newspaper reports tend to produce small networks with dramatic connections often to famous people. Court records produce discrete and highly structured small networks that make for good cases. If the OC [organized crime, ck] people from the FBI are involved the networks will almost inevitably look like 5 hierarchical families. If you

get less processed data (such as phone calls or surveillance reports), the world will look messier and less structured. Almost everyone will end up in one big component (which is of course interesting) but it will look more like the regular crime networks.”

I would add that anyone familiar with surveillance reports knows that they can be inaccurate, involve hearsay, and may reflect the biases of informants as well as those of the investigators.

3.2. *Second parties*

Network data readily include names of persons who were not respondents but who were named by the respondents. There is no convenient way of gaining their consent for the inclusion in the study unless the design involves a “snowball” (Thompson, 2002) in which these persons are also contacted for an interview. The procedures invoked for the original respondents in terms of confidentiality must be applied to these second parties, and in any case, if respondents are to cooperate, they need to be reassured that such safeguards apply to anyone they name. In the case of network snowballs including “multiplicity” sampling (Sirken, 1970; Sirken, 1998) one must ask the original respondent if the person named may be contacted and whether the original respondent’s name can be used as having given the referral. Respondents can be motivated to help if they are told that the person they referred will be given an incentive in their name, for example, “Joe sent me and wants us to give you a token honorarium of US\$ 25.” In our studies of Vietnam Veterans, respondents were generally happy to be in the position of make a “gift” to a veteran child, sibling, or parent (Boulanger and Kadushin, 1986; U. S. House of Representatives Committee on Veteran Affairs, 1981). Some IRBs balk at any incentives, however, despite their use in many U.S. government sponsored surveys.

More complicated is the use of names in studies that utilize publicly available data or databases. One might think that because the data are publicly available there is no ethical issue in the use of either individual or organization names. But the purpose of any statistical analysis, including network analysis, is to make visible that which is not apparent to the “naked eye.” Ronald Breiger (personal communication) offered the example of a regression analysis of posted faculty salaries that are available for many public universities. If the analysis includes the names of outliers, those whose earnings are two standard deviations below or above the expected values for persons with the same qualifications, has the privacy of those individuals been violated? The data are “legally” available, but the individuals whose names are used probably would not have the expectation that their salary status would be so compared with their colleagues. A book review of my *American Intellectual Elite* (Kadushin, 1974) complained that one of the reviewer’s friends, whom he named, was *not* included in my list of leading intellectuals. As I said, it may have been a mistake from many points of view to publish the list, but the privacy rights of the individuals named was certainly one of them. Though the intellectuals were happy to be named as “leading,” I had not considered the case of those who felt they were left out⁸.

⁸ Contrary to the impression of many reviewers who were not social scientists, the list was not “my” list: it was composed of persons most often nominated by the respondents.

3.3. Institutional review boards

The topic that raised the greatest concern at the Sunbelt Meetings, and has also continued to generate posting on the Social Network List Serve, is the reluctance of IRBs to approve network data collection. Problems with IRBs are frequent with all survey research, not only with network analysis, as the large volume of discussion on this topic on the American Association of Public Opinion Research (AAPOR) List Serve suggests. Even though most surveys can in principle be exempted from IRB detailed reviews, many IRBs are dominated by biological and medical researchers who engage in experiments and have little familiarity with large-scale efforts that deal with essentially anonymous respondents. They tend to reject survey protocols out of hand even though names of respondents are carefully guarded.⁹ AAPOR members have offered many useful suggestions, but since each IRB can enforce its own guidelines as long as they are stricter than the minimum required by the Federal Government, in the end each IRB must be individually persuaded that surveys do not violate rights to privacy. Network studies that are in principle not anonymous, even if only the researchers ever see the names, face even greater difficulties. INSNA formed a committee to try to offer a general set of guidelines to IRBs so that they might be better informed because the general sense of the session was that many IRBs do not understand the issues and so tend to reject any requests for the use of names. This has become a special problem in school classroom studies that used to be one of the easiest venues for network analysis. Data collection with minors as subjects has special rules, especially when drug and alcohol consumption, not to mention sexual activity is involved. Positive parental consent is generally now required. That is, instead of sending in a note that they object to their child's being included in a study, they now must send in a note specifically granting their child's participation. Since children are notorious in not transmitting notes from the school to their parents or guardians, many children can be omitted from research merely through administrative inefficiency. While many schools try to get blanket consent at the beginning of the school year and this may be satisfactory for some research, in social network data collection the systematic absence of some nodes can cause serious interpretation problems. Some schools do engage in a strong attempt to contact parents, but some parents feel safer in the default position of not giving consent. The result is that some students are singled out: when most of the class goes to the computer lab to participate in the network study non-participants are quite visible.

But IRBs can also be helpful, as long as they have approved the network protocol. Joe Labianca, School of Business, Emory University writing also on behalf of his colleague Jonathan L. Johnson of Sam M. Walton College of Business reported some aspects of their workshop at the 2000 Meetings of the Association of Management. The workshop was about how to get people to give negative network data, e.g. who they *don't* like, so the problem of eliciting responses was especially acute. They had the advantage, useful in this

⁹ The names cannot be immediately discarded for several reasons: first, to insure higher response rates sampled respondents are recontacted and many as 20 attempts to find them are typical of very high quality academic research; second, a proportion of respondents are recontacted by a supervisor to insure that interviewers did not invent the responses out of whole cloth; finally, panel studies require that names and addresses be kept in some secure form.

case, of not being consultants. They begin by conducting introductory meetings with all employees explaining the purpose of the study and its ultimate utility to mankind; they stress confidentiality issues and point out that the study has been approved under federal guidelines governing human subjects. In additional communication to me they noted that with the approval of their IRB obtained, one “can apply for a Certificate of Confidentiality from the National Institute of Health (NIH). This protects the network data from being subpoenaed for a civil or criminal suit.” To show how tricky this can get however, a recent experiment suggested that mention of this certificate could *increase* perception that harms might ensue (Singer, 2004).

3.4. *In conclusion: who benefits*

The issue of who benefits from sociometric research underlies this entire discussion. We take it for granted that we should do no harm. Though what constitutes harm can be a matter of serious controversy, I find it easier to address the benefit issue since if the benefits are substantial, they will usually outweigh whatever minimal harm is engendered by carefully managed network data and analysis. Benefits are a tricky subject. Labianca and Johnson attempt to persuade their subjects that what they are doing will be of benefit to humanity. This is a general assumption of science and especially invoked in medical research. Even though the individual subject may not benefit from the research, lives may eventually be saved. This is especially evident in epidemiology. But the truth is that neither science nor humanity in the abstract is the beneficiary in academic research. Rather, the investigators are the beneficiaries: their ego, prestige, science citation index and/or salaries are at stake. Those who teach may gain the added advantage of a live example for their classes. This is not to be deplored for it underlies the institutional structure of science. I do not necessarily advocate changing our more abstract appeal to subjects to note this reality (though graduate students and Ph.D. candidates can be effective in directly appealing to subjects to help the students pass their course requirement or to help them get a Ph.D.), but it is worth putting the matter of benefits into its proper context. *We* are the prime beneficiaries.

Then there are benefits to particular collectivities. In organization research on behalf of clients, the organization benefits, if the research fulfills its objectives. To the extent that organizational benefits are shared with employees – it becomes a better place to work, life is made more rational or easier, or everyone makes more money – this is fine. If the benefits accrue only to management or worse, to particular managers, then the researcher might carefully consider the ethics of the situation. Network research into criminals or terrorists is supposed to benefit the nation or the state. Here it is clear that the individual subjects not only do not benefit, but also are intentionally put in harms way. As noted above, there are two ethical issues at stake: the first is whether or not network “fingering” will be subject to judicial review, and second, the extent to which the data are reliable with the accuracy necessary when people’s lives are at stake—either the individual’s or the collectivities. The first issue is not discussed in this paper, though many will have strong opinions on it; the second was discussed and my opinion is that the data are rarely of sufficient quality for life and death decisions, or even for purposes of incarceration. The two issues are linked, however. Open court judicial review in an adversarial procedure could assess the accuracy

of the network data, as happens with other data in criminal procedures, but such reviews are rarely invoked, or invoked after accusations have been publicly leaked.

Benefits to second parties are critical in epidemiological investigations. Second parties can become aware that they are at risk for a disease, though at some cost to respondents who may be “outed,” if this concept can be applied to any carrier of a disorder or possessor of a socially undesirable characteristic. On the other hand, the benefits to society at large of containing disease are enormous.

Finally, there are direct benefits to individual respondents, though I believe these situations are rare. Moreno’s original situation with the girls in the cottages was “win-win.” The girls could live with people they liked, and the general obstreperousness of the cottages was reduced. But the second benefit was one that accrued to the management of the institution and so belongs to the category of benefit to the collectivity. Most gains for individual respondents are distributed gains that come about as a result of a gain to the collectivity. In network sampling, however, respondents directly gain the credit of having given relatives a gift.

The reader will have noticed that I have raised more questions than I have answered. I think this is a requirement of discussions of ethics.

References

- Alba, R., Guttman, M., 1972. SOCK: a sociometric analysis system. *Behavioral Science* 17, 326.
- Alba, R.D., 1973. A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology* 3, 113–126.
- Alba, R.D., Kadushin, C., 1976. The intersection of social circles: a new measure of social proximity in networks. *Sociological Methods & Research* 5 (1), 77–102.
- Alba, R.D., Moore, G., 1978. Elite social circles. *Sociological Methods & Research* 7 (November), 167–188.
- Borgatti, S.P., Everett, M.G., Freeman, L.C., 2004. *Ucinet 6 for Windows Computer Software*. Analytic Technologies, Harvard.
- Borgatti, S.P., Molina, J.L., 2003. Ethical and strategic issues in organizational social network analysis. *The Journal of Applied Behavioral Science* 39 (3), 337–350.
- Boulanger, G., Kadushin, C. (Eds.), 1986. *The Vietnam veteran redefined: fact and fiction*. Lawrence Erlbaum, Hillsdale, NJ.
- Fasshi, F., 2003. December 18. Two novice gumshoes chartered the capture of Saddam Hussein. *Wall Street Journal*, 1. http://online.wsj.com/article_email/0,SB107170894237083800-H9jeoNolaN2opymbICHa62Cm4,00.html.
- Finckenaue, J., Waring, E., 1997. Russian Emigre crime in the U.S.: organized crime or crime that is organized? In: Willians, P. (Ed.), *Russian Organized Crime: The New Threat?* Frank Cass, London.
- Finckenaue, J., Waring, E., 1998. *Russian Mafia in America: Crime, Immigration, and Culture*. Northeastern University Press, Boston.
- Frank, K.A., Yasumoto, J.Y., 1998. Linking action to social structure within a system: social capital within and between subgroups. *American Journal of Sociology* 104 (3), 642–686.
- Gertner, J. 2003. December 14. The 3rd annual year in ideas. *The New York Times Magazine*, Late Edition, Final, Section 6, 92.
- Higley, J., Hoffmann-Lange, U., Kadushin, C., Moore, G., 1991. Elite integration in stable democracies: a reconsideration. *European Sociological Review* 7 (1), 35–53.
- Higley, J., Moore, G., 1981. Elite integration in the United States and Australia. *American Political Science Review* 75 (3), 581–597.

- Kadushin, C., 1969. The professional self-concept of music students. *American Journal of Sociology* 75 (3), 389–404.
- Kadushin, C., 1974. *The American Intellectual Elite*. Little Brown, Boston, MA.
- Kadushin, C., 1995. Friendship among the French financial elite. *American Sociological Review* 60 (2), 202–221.
- Kadushin, C., Abrams, P., 1993. Social structure of Yugoslav opinion-makers: part I, informal leadership. In: Barton, A., Denitch, B., Kadushin, C. (Eds.), *Opinion-Making Elites in Yugoslavia*. Praeger, New York, pp. 155–187.
- Kadushin, C., Abrams, P., 1993. Social structure of Yugoslav opinion-makers: part II, formal and informal influences and their consequences for opinion. In: Barton, A., Denitch, B., Kadushin, C. (Eds.), *Opinion-Making Elites in Yugoslavia*. Praeger, New York, pp. 188–219.
- Klov Dahl, A.S., 2005. Social network research and human subjects protection: towards more effective infectious disease control. *Social Networks* 27 (2), 119–137.
- Krebs, V.E., 2004. Unclouthing terrorist networks. http://www.firstmonday.org/issues/issue7_4/krebs/.
- Krebs, V.E., 2003. Orgnet.com. <http://www.orgnet.com/>.
- Merton, R.K., 1968. *Social Theory and Social Structure*. Free Press, Glencoe, IL.
- Metz, C., 2004. January 20. Make contact. *PC Magazine*, 131–138.
- Mintz, B., Schwartz, M., 1985. *The power structure of American business*. University of Chicago Press, Chicago, IL.
- Mizruchi, M., 1996. What do interlocks do? An analysis, critique, and assessment of research on interlocking directorates. *Annual Review of Sociology*, 271–298.
- Moore, G., 1979. The structure of a national elite network. *American Sociological Review* 44, 673–692.
- Moore, J., 1961. Patterns of women's participation in voluntary associations. *American Journal of Sociology* 66, 592–598.
- Moreno, J.L., 1953. *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama*. Beacon House, New York.
- Pentland, A., 2004. "Reality Mining" the Organization. http://www.technologyreview.com/articles/wo_pentland033104.asp?trk=n.
- Roethlisberger, F.J., Dickson, W.J., 1939. *Management and the Worker*. Harvard University Press, Cambridge, MA.
- Shelley, G.A., Bernard, H.R., Killworth, P.D., Johnsen, E.C., McCarty, C., 1995. Who knows your HIV status? What HIV+ patients and their network members know about each other. *Social Networks* 17, 189–217.
- Singer, E., 2004. Confidentiality assurances and survey participation: are some requests for information perceived as more harmful than others? In: Cohen, S.B., Lepkowski, J.M. (Eds.), *Eighth Conference on Health Survey Research Methods*. National Center for Health Statistics, Hyattsville, MD.
- Sirken, M.G., 1970. Household surveys with multiplicity. *Journal of the American Statistical Association* 63, 257–266.
- Sirken, M.G., 1998. Network sampling. In: Armitage, P., Colton, T. (Eds.), *Encyclopedia of Biostatistics*. John Wiley, Chichester, New York, pp. 2977–2986.
- Teten, D., Fisher, D., Allen, S., 2004. The Five Keys to Building Business Relationships Online. *On Line Business Networks* <http://www.onlinebusinessnetworks.com/books.htm>.
- Thompson, S.K., 2002. *Sampling*, 2nd ed. John Wiley, New York.
- U.S. House of Representatives Committee on Veteran Affairs, 1981. *Legacies of Vietnam: Comparative Adjustment of Vietnam Veterans and their Peers*. U.S. Government Printing Office, Washington, D.C.
- Waring, E., 2002. Conceptualizing co-offending: a network form of organization. In: Waring, E., Weisburd, D. (Eds.), *Crime and Social Organization: Advances in Criminological Theory*. Transaction Publishers, Rutgers, NJ.
- Wellman, B., 2003. Ties & bonds. *Connections* 25 (2).