Available online at www.sciencedirect.com
ScienceDirect
Social Networks 29 (2007) 44–58
SOCIAL NETWORKS
www.elsevier.com/locate/socnet

To tell the truth: Measuring concordance in multiply reported network data☆

jimi adams *, James Moody

Ohio State University, 300 Bricker Hall, 190 N Oval Mall, Columbus, OH 43201, USA

Abstract

Social network data must accurately reflect actors' relationships to properly estimate network features. Here, we examine multiple reports of sexual, drug-sharing and social tie data on high-risk networks in Colorado Springs. By comparing multiple reports on the same ties, we can evaluate the reliability of this study's network data. Our findings suggest that these data have a high level of reporting agreement. From these findings, we discuss implications for analysis of these and similar data and provide suggestions for future social network data collection efforts.
© 2005 Elsevier B.V. All rights reserved.

Keywords: Methodology; Data collection; Link-tracing; Reliability

1. Introduction

Network analysts are often interested in the potential flow of goods between nodes (Potterat et al., 2002; Woodhouse et al., 1994; Moody, 2002). Network measures, however, assume accurate data on the ties connecting actors (Liljeros et al., 2003; Marsden, 1990) and any misreported relationships can potentially result in biased estimates of relevant statistics (Butts, 2003; Rothenberg et al., 1995). For example, linking claims about connectivity in sex and needle-sharing networks to epidemic potential requires reliably reported sex and needle partnerships (Brewer et al., 1999, 2006). Unfortunately, data limitations make reliability tests impossible for most studies, and

☆ This work is supported by NIH grants DA12831 and HD41877. We would like to thank Steve Muth and Martina Morris for continued support in access to and understanding of these data. We also thank John Potterat, David C. Bell, Devon Brewer, the Social Structure Research Group at Ohio State University, and anonymous reviewers from *Social Networks* for helpful comments on previous drafts of this paper. Any mistakes remaining are the sole responsibility of the author(s). Please direct any correspondence to jimi adams, Department of Sociology, 300 Bricker Hall, 190 N. Oval Mall, Columbus, OH 43210 (adams.644@sociology.osu.edu).
* Corresponding author.
E-mail address: adams.644@sociology.osu.edu (j. adams).

0378-8733/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.socnet.2005.11.009

researchers are forced to take such data at face value. With complex linked ego-network designs, this means that researchers are often forced to use single reports of individuals about their own behaviors *and* the behaviors of those to whom they are connected (Moody et al., 2005; Rothenberg et al., 2001; Burt, 1985). When one has multiple reports of the same tie, however, there is a unique opportunity to examine network reliability.

In this paper, we examine the data reliability in the Colorado Springs, Project 90 data. We compare agreement across multiple reports of the sexual, drug-sharing and social ties for reliability. We examine the consistency of an individual's reports of the same information over time, their agreement with others when reporting ties of which they are members, and when reporting ties in which they are not involved. Overall, the results from these data suggest remarkable reporting reliability. These results suggest general confidence in high-risk activity data and have specific implications for the data collection approach used to generate the Project 90 data that could prove useful for future studies (Moody et al., 2005; Rothenberg et al., 2001; Friedman et al., 1997).

2. Background

Past research has evaluated reliability for network reports for both general network data (e.g., McCarty et al., 1997; Marsden, 1993; Burt, 1986) and sexual and drug user data specifically (Bell et al., 2000; Brewer et al., 1999). The levels of reliability vary greatly across these studies and have not yet yielded generally accepted reliability standards.

Marsden (1990) summarized many of the previous evaluation studies, finding 40–60% accuracy of communication network reporting in the now classic Bernard, Killworth and Sailer studies (e.g., Bernard and Killworth, 1977; Killworth and Bernard, 1976; Bernard et al., 1980). Freeman et al. (1987) find that respondents are accurate in reporting other people's seminar co-participation only 48% of the time. Work to that point however, largely attempted to evaluate studies that focus on relationships like communication and friendship, which are of a different quality than those evaluated here. Since then, Cascairo et al. (1999) found only 41–45% of friendship tie nominations are reciprocated, a comparison that is similar to the social-contact comparisons we will explore in this paper. Bell et al. (2000) evaluated the reliability of drug-sharing partnerships and found 40–88% agreement across multiple reports, which we will compare to the drug-sharing relationships in Project 90 data. For sex related behaviors, the comparisons are frequently limited to marital partnerships (e.g., Bignami-Van Assche and Watkins, 2004), and are therefore not directly comparable to the relationship reports evaluated here, although agreement for use of family planning methods is around 75% (see Table 3, Bignami-Van Assche and Watkins 2004:26). Additionally, others have estimated reporting reciprocity using "consensus networks" that combine multiple reports (see Batchelder et al., 1997). Finally, Brewer et al. (2006) have shown how agreement across reports, particularly with respect to relationship timing can have significant effects on our ability to measure concurrency. Others are developing methods for incorporating these estimates of data quality (or conversely, the error therein) into modeling strategies (e.g., Butts, 2003), and better estimates of existing data reliability can improve these efforts. However, most researchers are currently forced by the data collection strategy to take each nominated relation as given.

When applied to sexual or needle sharing networks, accurate network representation has direct implications for estimating STI transmission (Potterat et al., 1999; Klovdahl, 1985). While previous studies have examined node-level sampling validity within Project 90 data (Rothenberg et al., 2001; Muth et al., 2000; Woodhouse et al., 1994; Brewer et al., 2006), the reliability of edge reports has not been examined. The unique data collection design, which provides multiple reports on the same relation, makes these data ideal for assessing reliability.

2.1. Project 90

Project 90 was a federally funded, US Centers for Disease Control (CDC) project focused on HIV transmission risk in heterosexual and injecting drug user (IDU) populations. The primary aim was to identify and interview as much of the target population as possible (IDU, prostitutes, and their sex partners), and assess the size, structure and epidemic potential of the high-risk partnership network. Data were collected from 595 respondents using face-to-face interviews over a 5-year period using an open cohort design. Detailed overviews of the study (Potterat et al., 2004; Muth et al., 2000; Darrow et al., 1999), the representativeness of the sample (Rothenberg et al., 2001; Woodhouse et al., 1994) and the sampling methods used (Morris, 1997; Klovdahl, 1999; Anderson et al., 1990; Watters and Biernacki, 1989; Biernacki and Waldorf, 1981) are available elsewhere.

Project 90 respondents were asked about basic HIV/STD risk related behaviors, daily activities, and health status. The questionnaire also contained a local network module (Morris, 1997) for collecting data on sexual, needle-sharing, drug-sharing ties and close personal contacts; close personal contact included sharing clothes, meals or place of residence. This network module allowed the respondents to name each of *their* sexual, drug and social contacts. The link-tracing design of this study in turn allowed the contacts nominated by two or more respondents to be recruited into the study. A separate "matrix" portion of the data collection instrument allowed respondents to report on ties *among* their contacts as well as between their contacts and others not directly connected to ego (though this was relatively rare).

One of the primary uses of these data examines the impact of network structure on disease transmission (Potterat et al., 1999; Rothenberg et al., 1998, 1995; Klovdahl et al., 1992). Such network analyses are typically generated using *any* nomination of a tie reported in the data, which is common practice for both Project 90 data (Moody et al., 2005; Rothenberg et al., 1995) and elsewhere (Rothenberg et al., 2000, 2001). If dissonant reports are common, however, this approach may misestimate the corresponding risk network. While some previous research has compared reported relations data with observed relations (e.g., Bernard and Killworth, 1977; Bernard et al., 1980), similar observational data are nearly impossible to get for private or illicit behaviors (such as sexual contact or needle sharing). Thus, instead of comparing reports to observations, we estimate reliability using multiple reports of the same relation, measuring concordance/discordance among reports.[1]

If multiple reports of the same relationship are found to be in agreement, the use of these data collection approaches should continue and greater confidence can be placed in similarly collected *singly* reported network data. Dissonant reports, however, would raise obvious questions about the reliability of the data both for those respondents and for singly reported data collected in a similar manner. However, even if the data are found to be largely discordant, researchers could use the reliability estimates to simulate or weight networks by giving edges a probability based on the observed reliability.

Since it is cumbersome to name all possible combinations of reporting categories, we use Fig. 1 to summarize the different types of comparisons we make. Respondents ("R" in Fig. 1) are people who were sampled and provided information on three relationally defined classes of people: themselves, contacts and associates. Contacts ("C" in Fig. 1) are persons with whom the

[1] We note here that concordant reports do not *necessarily* imply a physical relation among the participants, since it is possible (though unlikely) for both respondents to falsely report that they were directly connected to each other. Due to the intensive research activity of the Project 90 team, while such outcomes are logically possible the probability is low.

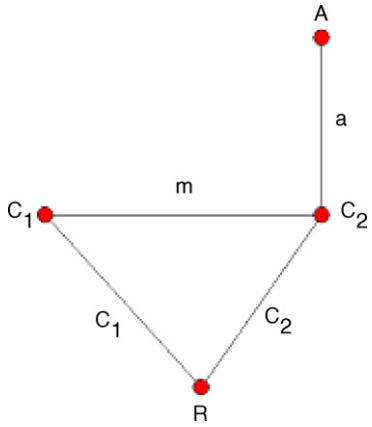

Fig. 1. Hypothetical four person sub-network.

respondent has a *direct* tie. Contacts may or may not also be respondents. Associates ("A" in Fig. 1) are a respondent's contacts' contacts, who are not directly tied to the respondent. Respondents were only asked to name up to one associate per contact, though interview timing often limited collection of any associate information. In reporting on relationships, there are also three dyadic classes: contact, matrix and associate ties. Contact ties ("c" in Fig. 1) are respondent reports of ties between that respondent and their contacts. Matrix ties ("m" in Fig. 1) are respondent reports of ties among their contacts. Associate ties ("a" in Fig. 1) are respondent reports of ties connecting their contacts to their associates. Fig. 1 presents a hypothetical four-person network, displaying all potential person and tie classes in relation to a single respondent (R). Project 90 data include information from 595 respondents about themselves and 8166 non-respondent contacts and associates. They report on 9953 contact ties, 23,977 matrix ties and 179 associate ties (pooled over time and across relation type), representing 29,477 unique dyads.

In addition to the information provided by R, persons C and A (Fig. 1) are also often respondents themselves, who can then report on the relations as seen from their own point of view, providing multiple observations of the same dyad. Reliable data would imply that any real relations with multiple observations would be multiply reported. However, while this comparison appears straight-forward, the *timing* structure of these data introduces more complexity. Because the Project 90 network instrument was time-specific, we need to include information on the timing of reports and interviews in selecting dyads that could have multiple reports.

2.2. Data structure

Relationship information was collected differently depending on the knowledge range respondents could reasonably be expected to provide. The investigators assumed that respondents would have more detailed information on their own relations than on relationships among their contacts. As such, we have detailed timing information on respondent ties but no specific timing information on matrix and associate ties. Specifically, respondents were asked if they had sex with, shared drugs with, or had social contact with each of their contacts *within the past 6 months*.[2] In the case

[2] Sex, drug and social contact ties in the data were actually recorded as *frequencies* of contact. For purposes here, they are further collapsed with any response of once or more in the past 6 months being coded as the presence of a tie, and a frequency of zero denoting the absence of a tie.

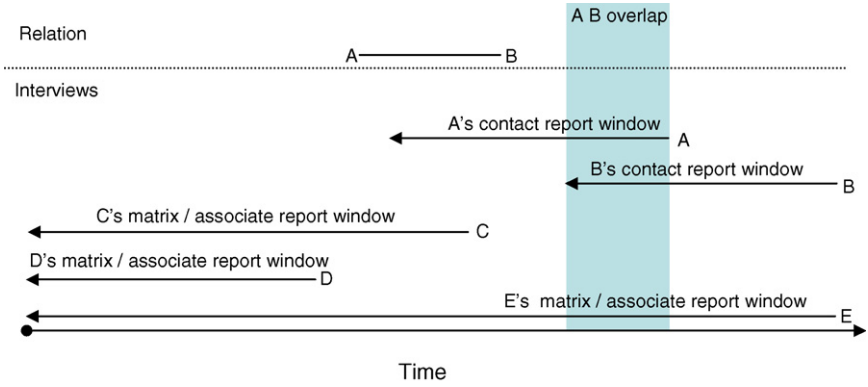

Fig. 2. Interview timing and relationship reporting.

of matrix and associate ties, such timed responses are not available as respondents were asked only if the two individuals were involved in each type of tie.[3] Respondents were also asked to distinguish the type of drug (needle or not) they shared with a partner, but they were not asked to distinguish among types of drugs for matrix and associate ties.[4] For the purposes of this analysis, with the exception of contact ties involving two respondents, drug-ties (whether needle-sharing or not) are collapsed into one drug tie category.[5]

Fig. 2 portrays a single relationship between two individuals and five potential respondents asked to report on that relationship. In the sections that follow, Fig. 2 helps clarify matches that are directly comparable, and which mismatches are potentially acceptable.

3. Methods

We first calculate frequencies and proportions of reporting concordance for each of the five comparisons described below by tie type. We make each comparison for social, sexual,[6] and drug-sharing ties. The universe for each of these comparisons differs considerably. Each of these comparisons is based upon different populations, and none of the groups examined are complete subsets of the others, even when they might intuitively seem so. Therefore, in the following sections we detail the question being asked in each of these five comparisons and describe the selection of the universe for each comparison. Following this description we will summarize the findings for each comparison.

[3] Interviewers did not follow an interview script per se, however, this question's basic structure was, "To your knowledge, do [person "a"] and [person "b"] KNOW each other/have SEX with each other/do DRUGS with each other (get high together)?" While the question is asked in the present tense, it is unclear what time referent the respondent is dealing with. As such, we treat such ties as "ever"—simply indicating that such a tie has been reported on that dyad.

[4] While not stated explicitly in the interview, investigators feel that most of the non-differentiated drug nominations involved needle exchange based on the type of drugs most of these respondents were using.

[5] This practice is also consistent with past studies (e.g., Moody et al., 2005; Potterat et al., 2004).

[6] In this paper, we do not include comparisons for any reports of same-sex sex ties. Because this data collection effort focused on heterosexual risk networks, we do not evaluate the potential for variance across interviews of the acceptance or elicitation of same-sex ties, because we are unsure of the effort used to collect ties that fell outside of the scope of the project. However, *including* such ties does not appreciably change the reporting concordance.

3.1. Contact report reliability

The simplest comparison is between reports of a relationship between two respondents. To examine reporting concordance, we select all dyads where a respondent ("R" in Fig. 1) nominates a contact tie (either "c_1" or "c_2"), and the contact ("C_1" or "C_2", respectively) is also interviewed. We expect that these reports should be highly concordant since respondents should have the best information about their own behavior.

As Fig. 2 makes clear, however, there are times where A and B could provide discordant information that is none-the-less reliable. If person B was interviewed more than 6 months after the end of the relationship, it would be possible for person A to report the relationship, person B to not report it, and both respondents to be providing reliable information. Thus, our second comparison selects all cases where a respondent nominates a contact *and* the contact is interviewed within the 6 months (before or after) reporting window. We expect that this timed comparison should provide even higher levels of reporting consistency than the pooled comparisons above since it more accurately reflects the interview questions.

3.2. Matrix/associate report reliability

Project 90 network data also include many reports of relationships provided by persons who are not a part of the relationship. Our expectation is that these nominations (matrix or associate) should be less reliable than contact reports.[7] We address three meaningful comparisons for matrix and associate tie nominations: reporting consistency, contact corroborations of matrix/associate reports, and matrix/associate corroborations of contact reports.

3.2.1. Reporting consistency

Respondents were interviewed up to five times over the course of the study, allowing us to examine their consistency of reporting the same relationship(s) over time. Since matrix and associate ties are not time-specific, our first examination of matrix/associate tie nominations simply compares the consistency of an individual's reports concerning the same tie across subsequent interviews. As can be seen in Fig. 2, as long as relationship AB happened prior to the interview date, E should provide the same information across any subsequent interview, if E had recent contact with both A and B at the time of subsequent interviews (or had contact with A or B and included B or A, respectively, among the named associates). We, therefore, examine the frequency that respondents report matrix/associate ties, given that they reported the tie existed in any previous interview. The comparison cases are therefore all interviews with "R" in years after the initial matrix/associate nomination.[8] We suspect that respondents provide relatively consistent information across time in this situation since their information should be consistent across time. Some previous research suggests that consistency may improve with subsequent reinterviews (Brewer, 2000), potentially indicating the role of increasing interviewer-subject trust or simply better recall over the course of longer-term research projects.

[7] To evaluate reports made about ties not including the respondent, all analyses in this paper include matrix and associate reports simultaneously. Of these reports, over 98% are matrix ties. All of the results presented here for matrix *and* associate ties do not differ significantly if the comparisons are limited only to matrix ties. In many of these comparisons the numbers of associate nominations that have data available for comparison are too small to warrant separate evaluation.

[8] This estimate is probably an under-count if respondents interpreted the question to mean "current" instead of "ever." If this were the case, then some of the "inconsistent" reports could be accurately capturing relations that ended.

3.2.2. Comparing matrix/associate reports to contact reports and each other

The final set to evaluate is the reliability of matrix/associate tie nominations. There are two ways to do this: first by comparing them to the contact reports evaluated above, and second by comparing the agreement across different respondent's evaluation of the same matrix/associate ties. First, we evaluate the corroboration of matrix/associate reports by contact reports. That is, given that a respondent reports a matrix/associate tie, we identify all cases where at least one of the contacts was *also* interviewed and calculate the frequency and proportion of agreement for each of these reports.[9] We can also reverse the question, and ask, given a contact report of a particular relation, how often do matrix/associate nominations corroborate the report?[10] Finally, we tabulate each matrix/associate nomination made for each unique dyad, and calculate the proportion of agreement between reporting pairs. This comparison, as those above, tabulates agreement by reporting pairs (e.g., from Fig. 1—comparing C's report of relationship AB to D's, as well as D's to E's *and* C's to E's, for all respondents including the AB pair). The result is a proportion of the time that two (or more) differing respondents agree about the presence (or absence) of a matrix/associate tie.

4. Concordance findings

4.1. Respondent–respondent contact ties

Table 1 presents the frequency of concordant reports for ties where *both* nodes are respondents for all tie types.[11] These comparisons are presented for the pooled data covering all time points as well as those that limit the comparison window to reports that fell within 6 months of each other. A dyad can be included in the top panel only once, but multiple times in the bottom panel if both members of the dyad were interviewed multiple times with overlapping reporting windows.

In the pooled comparisons, the results (top panel, Table 1) indicate whether the two respondents *ever* agreed on a positive tie nomination. The bottom panel of Table 1 shows the timed comparisons for each unique dyad-date combination where the two respondents' interview dates are within a 6-month window.[12] In the bottom panel, it is therefore possible for the same tie to be multiply reported, in different interview periods, and be correctly included as a comparison case. It is important to understand that because of the different sorting for the calculations in Table 1, neither of the panels are a complete subset of the other. The "Timed Proportion" row of Table 1 provides the correct subsample comparison, indicating the proportion of all pooled matches that are uniquely identified timed matches.

[9] The pairs included in these comparisons are selected from the pool of individuals reporting on the tie, not on the dyad. As a result, some dyads can potentially be included in this comparison twice.

[10] There are also acceptable mismatches of these tie reports (e.g., if person D and person A in Fig. 2 are reporting on relationship AB), however, those differences are difficult to interpret. For simplicity, and to more readily compare the levels of agreement here to those presented above for contact reports, we do exclude time-related acceptable mismatches from the analyses in Tables 3 and 4. Further, their exclusion from this analysis did not significantly alter the proportions observed for these comparisons; it only reduced the number of cases included for comparison.

[11] By sorting uniquely by dyad rather than respondent, this disallows multiple counts of the same nomination. That is, if A (Fig. 1) nominates B and B nominates A, only one of these is initially selected. If, instead we sort on respondent, the comparison data would redundantly include the initial data. That is, B's nomination of A would occur in both the initial nomination pool as well as the comparison group.

[12] In interpreting the bottom panel of Table 1, readers should recognize that the period of overlap between the two interviews can range from 1 day (when interviewed six months apart) to 6 months (for those interviews conducted on the same day).

Table 1
Concordance frequencies for contact tie reports

	Sex tie	Drug tie	Needle tie[a]	Social tie
Pooled				
Match	174	203	36	274
Mismatch	32	81	53	75
Proportion matching	0.845	0.715	0.405	0.785
N unique dyads	206	284	89	349
N unique nodes/respondents in dyads	184	214	106	263
Timed proportion[b]	0.799	0.754	1.00	0.803
Timed[c]				
Match	184	198	61	286
Mismatch	13	43	26	37
Proportion matching	0.934	0.822	0.701	0.885
N unique dyads	115	171	75	209
N unique nodes/respondents in dyads	184	214	106	263

[a] Needle ties are also included in the drug tie category.
[b] The proportion of all pooled matches that are actually uniquely identified timed matches.
[c] These comparisons eliminate known "acceptable" mismatches—those with no window of overlap.

The results from Table 1 show remarkable agreement levels. First, ignoring the size of the overlap window, when a respondent reports a tie, 74% (regardless of tie type) of those contacts who were interviewed corroborate the tie report.[13] For sex ties, 85% of all positive nominations are ever corroborated. The comparable rates of agreement are 72% for drug ties, and 79% for social ties. These figures are consistently lower than the time period matches, however the reliability ranking across tie-types is the same. Second, looking at those interviewed within 6 months of each other (the second panel of Table 1, which is the relevant population, given the interview question) 86% of all time-specific positive tie nominations (regardless of tie type) are reciprocated. The results by tie-type indicate that 93% of all sexual ties, 82% of drug-sharing relationships (70% if constrained to needle sharing relationships), and 89% of all social ties are confirmed by both parties.

The "Timed Proportions" show that the vast majority of concordant contact reports in these data are time-specific matches (88%, regardless of tie-type) and not pooled matches. By tie type the comparable figures show that 80% of concordant sex tie reports, 75% of concordant drug tie reports (100% for needle ties), and 80% of all corroborated social ties are corroborrated within the 6-month reporting window. It therefore appears that individuals gave reliable reports of their ties *during the prior 6 months* about which they were asked. One implication of this finding is that the data collected are more accurate with respect to time than we might typically assume, and thus would allow researchers to be more specific in constructing network images from the data (Brewer et al., 2006). With some careful consideration of the shifting time windows implicitly included in these data, researchers could map time-specific networks. While such maps may be computationally more difficult, they should more accurately represent topics of interest, such as epidemic potential (Moody, 2002).

[13] In other words, a reciprocation of the nomination *at any point in time* (not necessarily within the 6 months window, as the question was posed).

Table 2
Frequency of recant of matrix and associate tie nominations

	Sex tie	Drug tie	Social tie
Total positive reports	1608	8866	23694
Initial positive reports	977	8328	21263
Subsequent positive reports	631	538	2431
Subsequent negative reports	31	85	67
Uniquely contributed	(25)	(71)	(60)
Proportion of subsequent reports that are inconsistent[a]	0.047	0.136	0.027
Uniquely contributed	(0.038)	(0.117)	(0.024)

[a] Calculated by dividing row 4 by (row 3 plus row 4).

4.2. Matrix/associate tie nominations

4.2.1. Matrix/associate internal consistency

For matrix/associate ties, the questions asked were not temporally restricted. Treating these as untimed relations, if a respondent ever lists a positive tie nomination, s/he should therefore also report that tie in all subsequent interviews, since knowing of a relation at time t_0 should imply knowing the relation at time t_1.[14]

Table 2 presents the frequency of report consistencies for matrix/associate ties by type of tie, for all interviews after a positive tie nomination.[15] The results suggest that inconsistent reports are infrequent, although drug ties have the lowest consistency. Conditional on a previous positive nomination, about 14% of subsequent reports deny a drug tie. Similar interpretations are possible for each of the other categories, and show that only 5% of sex tie reports and 3% of social tie reports subsequently made by matrix/associate nominations are inconsistent. Because of the different reasons for potential discrepancies, we are not able to determine which report is incorrect, but each of the mismatches shown in Table 2 indicates some level of misinformation or reports focused solely on current behavior.

4.2.2. Contact corroborations of matrix/associate nominations

The previous section indicates that individuals in the Project 90 data are remarkably reliable in the reporting of their own relationships. We, therefore, use contact ties to examine the reliability of matrix/associate ties. Since there is no reason for researchers to assume that respondents would be more likely to mis-represent information for relationships that cannot be similarly evaluated, the results of this comparison can then also presumably be extended to provide evidence of the reliability of matrix/associate nominations in general. Our suspicion is that these nominations may be less reliable than contact ties, although past research has treated them as equally reliable in constructing the risk network (Rothenberg et al., 1998; Darrow et al., 1999).

Table 3 presents the proportion of matrix/associate reports that are corroborated by contact reports of the same tie. That is, given that C_1 or C_2 (Fig. 1) reports a tie between AB, how

[14] Again, these will be conservative estimates to the extent that respondents treated these as "current" instead of "ever" ties.

[15] This may underestimate reporting inconsistencies. It is possible in subsequent iterations that a respondent providing information on matrix associate ties could: (1) rename the same tie in a later interview; (2) include C_1 and C_2 among their subsequent lists of alters, but fail to report the same tie between C_1 and C_2; (3) or fail to include C_1 and/*or* C_2 in their subsequent lists of alters, and therefore not provide any information about that tie in subsequent years. The estimates calculated here do not include the potentially inconsistent reports represented by (3). Thanks to David Bell for clarifying this point.

Table 3
Contact corroborations of matrix and associate tie reports

	Sex tie	Drug tie	Social tie
Match	126	573	709
Mismatch	30	116	351
Proportion matching	0.808	0.832	0.669
N unique dyads	85	347	597
N unique nodes in dyads	135	268	422
N unique respondents	166	297	450

Table 4
Matrix and associate corroborations of contact tie reports

	Sex tie	Drug tie	Social tie
Match	117	536	726
Mismatch	91	182	111
Proportion matching	0.563	0.747	0.867
N unique dyads	130	384	486
N unique nodes in dyads	160	259	378
N unique respondents	207	309	428

frequently does A or B also report that tie. It is important to note that this comparison sorts on unique respondent–contact combinations, and therefore does not multiply include any dyad-respondent pairs over time, but does potentially allow for multiple comparisons of the same dyad from multiple respondents. Unlike the findings above for contact ties, matrix/associate sexual tie nominations are *not* the most likely to be corroborated. This makes sense, however, because sexual behavior is less likely to take place in the presence of others, so accurate information is harder to attain for third parties. As expected, the level of agreement across relations tends to be lower than that for contact tie agreement. The information in Table 3 tells us that 81% of all matrix/associate sex tie nominations are corroborated by contact reports of the same tie, and the agreement is 83% for drug ties and 67% for social ties.[16] The overall lower percentage of agreement here suggests that matrix/associate tie nominations are slightly less reliable than the levels of agreement for contact ties, though still quite high in comparison to levels that might be expected based on prior work. The figures for drug ties are actually remarkably similar to (even slightly higher than) the levels found for contact comparisons. The data here appear to suggest that drug behavior is relatively well known to contacts and acquaintances, while sexual behavior is slightly less well known, and awareness of social interactions between individuals may simply be over-assumed on the part of respondents.

4.2.3. Matrix and associate corroborations of contact tie reports

Table 4 presents the reverse question, showing the proportion of all positive self-reported contact ties that are corroborated by other respondents' matrix/associate nominations of the same tie. As above, sex ties are the least likely reports to be corroborated, with approximately 57% of matrix/associate responses effectively identifying a sex tie that was reported by those involved in

[16] Calculated as $\left(\sum_i \mathrm{P}(\mathrm{C}_{(\mathrm{m})}|\mathrm{R}_{(\mathrm{m})}) + \sum_i \mathrm{P}(\mathrm{C}, \mathrm{A}_{(\mathrm{a})}|\mathrm{R}_{(\mathrm{a})})\right) / n_{\mathrm{R}(\mathrm{m},\mathrm{a})}$, for all m and a.

Table 5
Matrix and associate nomination agreement

		Sex tie	Drug tie	Social tie
Summed by dyad				
Proportion agree		0.898 (0.283)	0.783 (0.384)	0.740 (0.411)
N unique dyads		2047	2094	2120
By *N* respondents per dyad				
2	Proportion agree	0.919 (0.273)	0.816 (0.388)	0.765 (0.424)
	N unique dyads	1777	1779	1772
3	Proportion agree	0.775 (0.316)	0.573 (0.321)	0.583 (0.323)
	N unique dyads	178	209	248
4+	Proportion agree	0.720 (0.281)	0.801 (0.266)	0.786 (0.268)
	N unique dyads	92	106	100

Note: numbers in parentheses are standard deviations.

the relationship. For drug and social tie nominations the levels of agreement approaches that for contact tie reports (75% agreement for drug ties and 87% agreement for social ties).

The final analysis, we present here is the levels of agreement observed among multiple matrix/associate reports of the same tie. For each dyad we selected all respondents who reported on that tie. We then compared each respondent–respondent pair to examine whether they agreed on their report of the tie (1) or not (0).[17] These levels of agreement were then averaged for each dyad by tie type. We then summarized those proportions across each of the three tie types. The first panel of Table 5 reports the levels of agreement for all reports, for all dyads, including the level of agreement for each dyad only once. That is, whether there are two matrix/associate nominations or three (or more), the level of agreement for each dyad contributes equally to the overall proportion. The second panel generates agreement levels conditional on the number of reports for each dyad.

Matrix/associate tie nominations agree 90% of the time for sex tie nominations, 78% of the time for drug-sharing relationships, and 74% of the time for social relationships. The results in the bottom panel of Table 5 suggest two additional insights. First, the vast majority of these comparisons (approximately 85% for all tie types) include only two matrix/associate reports. These figures show that 92% of doubly reported matrix/associate sex tie reports agree, while those reported by three persons agree 78% of the time, and those by four or more agree 72% of the time.[18] The other observation of note here is that for social ties and drug-sharing relationships, those most reported are the most likely to agree, which is consistent with the fact that these are potentially more public behaviors than are sexual relationships.[19]

[17] These comparisons are slightly different than those reported above, as they are not conditional on a positive report. In these comparisons, agreement is considered for all respondent–dyad pairs, not just those including at least one positive report.

[18] Table 5 warrants one further note of interpretation. For the bottom panel, the potential values are constrained. When there are only two reports, agreement is either 1 or 0. For those including three reports the potential values are limited to 0.333 or 1. For those including four reports, the potential values are limited to 0.33, 0.5 and 1.

[19] We also calculated the correlation between the levels of agreement observed between matrix/associate nominations (presented in Table 5) and those observed between contact reports (presented in Table 1). Those correlations were positive for all tie types, but because of issues of case selection incompatibility and constrained potential values for the matrix/associate comparisons (see Note 18), these correlations did not achieve standard levels of significance.

We also examined the possibility of any significant predictors of report inconsistency for matrix/associate tie nominations using logistic regression. Those analyses suggest that while matrix/associate ties, as a group, may be slightly less reliable than contact reports, there does not appear to be any systematic differences (whether characteristics of the respondent or the dyad) in their reporting consistency that would preclude their use in constructing other network measures. Results of these additional analyses are available from the first author upon request.

5. Discussion

5.1. Reliability of contact tie nominations

For the most part, when two individuals report on contact with each other they agree on the nature of the tie. This suggests that Project 90 respondents provided reliable information for contact ties. While needle-sharing ties are slightly less reliable than the other types of ties, all are significantly better than previous literature would suggest. The higher rates of agreement for sex ties than for drug or needle sharing relationships may indicate forgetting of such ties, which has been found in other research (Brewer et al., 1999). Moreover, respondents appear reliable for both who they are connected to and when they were connected (Brewer et al., 2006). Since our constructed overlap can range between 1 day and 6 months, it is likely that the respondents are using a "fuzzy" 6-month window, though they appear to be doing so consistently. It clearly does not appear that respondents answer time-specific questions with time-free responses. Since relationship timing is crucial for disease diffusion (Moody, 2002; Morris and Kretzchmar, 1995) these findings suggest that respondents can provide useful information on relationship timing.

5.2. Reliability of associate and matrix tie nominations

People are not as reliable reporting non-contact ties as they are for their own ties. While the proportions across tables are not directly comparable (because of different selection criteria), there is a clear decline in reporting concordance. In these data, respondents appear to slightly underreport matrix and associate ties. While prior work has treated non-contact ties similarly to contact ties, the lower reliability suggests treating them as 'probable' instead of 'actual' and adjusting networks accordingly. The effect of this change will, of course, depend on the structure of the overall pattern of non-contact ties, but one could run sensitivity analyses by re-calculating network statistics with varying proportions of non-contact ties retained in the network. The best way to do this would include information on dyad types from the multiply reported dyads.

The levels of agreement that we observe in the Project 90 data are clear improvements over levels that are reported in most previous studies. For example, in the estimates here, reports are consistent for contact ties 70–93% of the time (depending on tie type), and matrix/associate ties are corroborated by 67–83% of contact tie nominations, and corroborate contact tie nominations 56–87% of contact tie nominations. For an alternate interpretation, we could think of the complements of the results in Table 3 as approximations of the false positive reports among matrix/associate tie nominations and those in Table 4 as approximating false negatives.

Most studies that examine reporting concordance for various relations between persons in a dyad show between 40 and 60% agreement (Marsden, 1990:448, and see review above). The only studies that show levels of agreement similar to Project 90 compare relationship reports that are somewhat more restrictive. Bell et al. (2000), for example, find 70–88% agreement in a study similar in structure to Project 90 (see Table 3), though their comparisons are based on reports

only covering the last 30 days. Combining our results with Bell et al. (2000) findings, suggests that accurate representation of relationships *and* relationship timing is possible with careful study design.

Agreement levels in non-contact reports of social ties may be additionally informative. Table 4 shows that 87% of the time a respondent reports a social tie between two of their contacts or associates, those individuals also report the same tie. In contrast, if those individuals first report a social tie between themselves, the respondent also reports that tie 67% of the time (Column III, Table 3). What this does not show is that for all interviews regarding non-contact ties, respondents reported that the individuals knew each other almost 98% of the time. In other words, it appears that respondents are assuming that since they know two people, those two people must in turn know each other.[20] Respondents do not, however, appear to make the same kinds of generalizations about sexual or drug ties (the comparable figures for sex and drug ties are only approximately 7 and 37%, respectively). Thus, we suspect that any over reporting in social ties is due to projecting one's own ties, but that people do not make such generalizations about those more intimate social relations with epidemiologic implications.

6. Conclusion/implications

Project 90 respondents were remarkably reliable when reporting about the extent and timing of their own relationships. This is particularly relevant for processes such as disease diffusion, as it allows one to construct timed networks which can more accurately represent the epidemic potential of the network (Moody, 2002; Morris and Kretzchmar, 1995). While current uses of such data often ignore time, we suggest that researchers incorporate the timed data to more fully match available data to processes of interest. In contrast, our results suggest that researchers might temper the use of non-contact reports, since these are comparatively less reliable (though still higher than many prior reports might lead one to expect). One principled approach to using such data might be to treat such ties as probable within the reliability limits, and then generate ranges of networks based on these likelihood's (which can be had most simply by simulating multiple instantiations of the network based on the edge reliabilities; see Butts, 2003 for an alternative approach).

Since network analysts are interested in studying processes such as disease propagation that are time dependent, where estimates demand accuracy, these findings should be encouraging. What accounts for the high level of concordance in these data and are there hints from the Project 90 collection that others could use? We think much of the success of this project rests with the trust built between researchers and subjects that was built through the long-term community embeddedness of this project. This particular study employed a research staff that had many years of ethnographic and epidemiologic contact with the individuals being studied. The remarkable levels of data agreement in this study may therefore be attributable to respondents' willingness to name individuals involved in private or illegal activities because of high levels of trust that were developed over this long contact. This suggests that researchers interested in obtaining accurate data in such contexts will need to invest the time and energy required to build trust and knowledge of the community. This likely means partnering with local agencies with a good on-the-ground reputation and employing careful methods to ensure confidentiality and safety for respondents. In addition to this kind of local embeddedness, researchers should also build data redundancy into

[20] Note that this does not imply a total network density of 98%, but instead that these nets are very clumpy around central nodes.

their designs. By designing studies where relations will be multiply reported, one can estimate the local reliability of the data and thus place more accurate bounds on empirical claims.

References

Anderson, R.M., Gupta, S., Ng, T.W., 1990. The significance of sexual partner contact networks for the transmission dynamics of HIV. Journal of Acquired Immune Deficiency Syndrome 3, 417–429.

Batchelder, W.H., Kumbasar, E., Boyd, J.P., 1997. Consensus analysis of three-way social network data. Journal of Mathematical Sociology 22, 29–58.

Bell, D.C., Montoya, I.D., Atkinson, J.S., 2000. Partner concordance in reports of joint risk behaviors. Journal of Acquired Immune Deficiency Syndromes 25, 173–181.

Bernard, H.R., Killworth, P.D., 1977. Informant accuracy in social network data II. Human Communications Research 4, 3–18.

Bernard, H.R., Killworth, P.D., Sailer, L., 1980. Informant accuracy in social network research IV: A comparison of clique-level structure in behavioral and cognitive data. Social Networks 2, 191–218.

Biernacki, P., Waldorf, D., 1981. Snowball sampling: problems and techniques in chain-referral sampling. Social Methods and Research 10, 141–163.

Bignami-Van Assche, S., Watkins, S.C., 2004. Husband–wife disagreement in rural Malawi: a longitudinal analysis. Paper presented at the Annual Meetings of the Population Association of America, April, Boston. Available for download from (http://paa2004.princeton.edu/download.asp?submissionId=40391).

Brewer, D.D., 2000. Forgetting in the recall-based elicitation of personal and social networks. Social Networks 22, 29–43.

Brewer, D.D., Garrett, S.B., Kulasingham, S., 1999. Forgetting as a cause of incomplete reporting of sexual and drug injection partners. Sexually Transmitted Diseases 26, 166–176.

Brewer, D.D., Rothenberg, R.B., Muth, S.Q., Roberts, J.M., Potterat, J.J., 2006. Agreement in reported sexual partnership dates and implications for measuring concurrency. Sexually Transmitted Diseases, in press.

Burt, R.S., 1985. Network Items and the General Social Survey, GSS Technical Report No. 53, Chicago: NORC, Fall, 1984. Published in Social Networks 6, 293–339.

Burt, R.S., 1986. A Note on Sociometric Order in the General Social Survey Network Data, GSS Technical Report No. 60, Chicago: NORC, 1985. Published in Social Networks 8, 149–174.

Butts, C.T., 2003. Network inference, error, and informant (in)accuracy: a Bayesian approach. Social Networks 25, 103–140.

Cascairo, T., Carley, K.M., Krackhardt, D., 1999. Positive affectivity and accuracy in social network perception. Motivation and Emotion 23, 285–306.

Darrow, W.W., Potterat, J.J., Rothenberg, R.B., et al., 1999. Using knowledge of social networks to prevent human immunodeficiency virus infections: the Colorado Springs study. Sociological Focus 32, 143–158.

Freeman, L.C., Romney, A.K., Freeman, S.C., 1987. Cognitive structure and informant accuracy. American Anthropologist 89, 310–325.

Friedman, S.R., Neaigus, A., Jose, B., et al., 1997. Sociometric risk networks and HIV risk. American Journal of Public Health 87, 1289–1296.

Killworth, P.D., Bernard, H.R., 1976. Informant accuracy in social network data. Human Organization 35, 269–296.

Klovdahl, A.S., 1985. Social networks and the spread of infectious diseases: the AIDS example. Social Science Medicine 21, 1203–1216.

Klovdahl, A.S., 1999. Urban social networks: some methodological problems and possibilities. In: Kochen, M. (Ed.), The Small World. Ablex, Norwood, NJ (Chapter 10).

Klovdahl, A.S., Potterat, J.J., Woodhouse, D.E., et al., 1992. HIV infection in an urban social network: a progress report. Bulletin De Methodologie Sociologique 36, 24–33.

Liljeros, F., Edling, C.R., Amaral, L.A.N., 2003. Sexual networks: implications for the transmission of sexually transmitted infections. Microbes and Infection 5, 189–196.

Marsden, P.V., 1993. The reliability of network density and composition measures. Social Networks 15, 399–421.

Marsden, P.V., 1990. Network data and measurement. Annual Review of Sociology 16, 435–463.

McCarty, C., Bernard, H.R., Killworth, P.D., et al., 1997. Eliciting representative samples of personal networks. Social Networks 19, 303–323.

Moody, J.W., 2002. The importance of relationship timing for diffusion: indirect connectivity and STD infection risk. Social Forces 81, 25–56.

Moody, J.W., adams, j., Muth, S.Q., Morris, M., 2005. Quantifying the benefits of link-tracing designs for partnership network studies. Unpublished manuscript.

Morris, M., 1997. Sexual Networks and HIV. AIDS 97: Year in Review 11 (Suppl. A), S209–S216.

Morris, M., Kretzchmar, M., 1995. Concurrent partnerships and transmission dynamics in networks. Social Networks 17, 299–318.

Muth, S.Q., Potterat, J.J., Rothenberg, R.B., 2000. Birds of a feather: using a rotational box plot to assess ascertainment bias. International Journal of Epidemiology 29, 899–904.

Potterat, J.J., Phillips-Plummer, L., Muth, S.Q., et al., 2002. Risk network structure in the early epidemic phase of HIV transmission in Colorado Springs. Sexually Transmitted Infections 78, 159–163.

Potterat, J.J., Rothenberg, R.B., Muth, S.Q., 1999. Network structural dynamics and infectious disease propagation. International Journal of STD & AIDS 10, 182–185.

Potterat, J.J., Woodhouse, D.E., Muth, S.Q., et al., 2004. Network dynamism: history and lessons of the Colorado Springs study. In: Morris, M. (Ed.), Network Epidemiology: A Handbook for Survey Design and Data Collection. Oxford University Press, New York, pp. 87–114.

Rothenberg, R., Baldwin, J., Trotter, R., Muth, S.Q., 2001. The risk environment for HIV transmission: results from the Atlanta and Flagstaff networks studies. Journal of Urban Health 18, 419–432.

Rothenberg, R., Potterat, J.J., Woodhouse, D.E., et al., 1998. Social network dynamics and HIV transmission. AIDS 12, 1529–1536.

Rothenberg, R.B., Long, D.M., Sterk, C.M., et al., 2000. The Atlanta urban networks study: a blueprint for endemic transmission. AIDS 14, 2191–2200.

Rothenberg, R.B., Potterat, J.J., Woodhouse, D.E., et al., 1995. Choosing a centrality measure: epidemiologic correlates in the Colorado Springs study of social networks. Social Networks: Special Edition on Social Networks and Infectious Disease: HIV/AIDS 17, 273–297.

Watters, J.K., Biernacki, P., 1989. Targeted sampling: options for the study of hidden populations. Social Problems 36, 416–430.

Woodhouse, D.E., Rothenberg, R.B., Potterat, J.J., et al., 1994. Mapping a social network of heterosexuals at high risk for HIV infection. AIDS 8, 1331–1336.