

A Modified Elicitation of Personal Networks Using Dynamic Visualization

Christopher McCarty¹

Bureau of Economic and Business Research, University of Florida, Gainesville, FL

Sama Govindaramanujam²

Bureau of Economic and Business Research, University of Florida, Gainesville, FL

Several algorithms and software packages have been developed for displaying the relationship between actors within a whole (sociocentric) network. These visualization packages use as input an adjacency matrix representing the relationship between actors, and have occasionally been applied to personal (egocentric) network data. Personal network adjacency matrices require respondents to report on all alter-alter ties. This is an enormous respondent burden when the number of alters goes much beyond 30. We report here on an effort to reduce that burden by having respondents build their own personal networks, interactively, on the Internet. In a study on smoking, 100 respondents (50 smokers and 50 non-smokers) listed 45 network alters and provided data on whether each of the 990 pairs of alters talked to each other. We used a program called EgoNet to collect these data. Fifty of the respondents (25 smokers and 25 nonsmokers) then completed a similar exercise over the Internet, using a visual interface, called EgoWeb. There are clear mode effects on personal network composition and structure.

BACKGROUND

Many advances have been made in the visualization of network data over the past decade. Until the past couple of years, virtually all network visualization packages were oriented toward whole (sociocentric) networks. These packages, such as PAJEK, NETDRAW, KRACKPLOT, NETVIZ (among others), typically provide a variety of visualization algorithms using an existing adjacency matrix as input. The result is a two or three dimensional representation of the links within the group, and often the ability to display attributes of network nodes (or actors) using size, shape, color or some combination of these.

¹ Chris McCarthy, Survey Director, University of Florida Survey Research Center, Bureau of Economic and Business Research, PO Box 117145, University of Florida, Gainesville, FL 32611-7145, chrism@bebr.ufl.edu, fax: (352) 392-4739.

² Sama Govindaramanujam, Graduate Research Assistant, University of Florida Survey Research Center, Bureau of Economic and Business Research, PO Box 117145, University of Florida, Gainesville, FL 32611-7145, samag@bebr.ufl.edu.

Unlike whole networks, egocentric networks are centered on a focal individual. In the field of social network analysis, egocentric analyses are often done of members of a whole network. Some network visualization packages allow the user to visualize the egocentric network of a member of that whole network. For example, NETDRAW has a module that allows the user to pick a node in a whole network and display ties to the nodes to which it is connected. In practice, for example, one may want to toggle between viewing the entire set of ties between children in a classroom, or the egocentric network of a single student.

The logical extreme of an egocentric network is the personal network. Unlike any of the prior examples, a personal network is an egocentric network existing within the whole network defined by the population of the world. In other words, personal networks are not constrained by a sub-structure, such as geographic or social space. Personal networks can vary between structurally cohesive networks that are compositionally homogenous in terms of member characteristics, to compositionally heterogeneous networks that exhibit extensive bridging and reach across geographic and social space. Unlike their constrained counterparts, egocentric networks include the influences of all the whole networks to which a respondent belongs, the effect of the overlap between those whole networks, and the potential to use these characteristics as explanatory or dependent variables.

Researchers who want to understand the effect of personal networks face an enormous data collection problem. One cannot know at any given time the names of all the members of the world, and there is no practical way to present this list to a respondent. One can only ask respondents to whom they are tied. There is, however, bias in the way respondents list alters. Brewer (2000) found that both close and weak ties could not be recalled by respondents in a free-listing task. Brewer and Webster (1999) found that forgetting in a free list of alters affected the estimated structural features of a whole network. This research suggests that names are not recalled randomly from respondent memory. Other research has focused on cueing mechanisms to enhance recall to correct for this bias (Brewer and Garrett 2001; McCarty et al. 1997; Brewer 1997).

Another solution to the problem of recall bias is to generate a sample of personal network alters so large that the bias towards strong ties or those with particular characteristics is minimized. Some research suggests that personal networks consisting of active ties (those contacted in the past two years) are roughly size 290 (McCarty, et al. 2000; Killworth, Bernard and McCarty 1984). McCarty (2002) had respondents list 60 alters in a study of personal network structure. A sample of 60 alters would account for nearly 20 percent of the personal network and would presumably minimize recall bias.

Most personal network studies are on a small number of alters and almost never consider personal network structure, instead relying on personal network composition for explanatory power. While explorations of personal network compositional variables (such as the percent of network alters that are family, women or who smoke) are useful, the analysis of personal network structural variables (such as closeness or betweenness centrality and the number of components) remains largely unexplored. Previous studies have limited their analyses mostly to network density (Latkin et al. 1998; Haines, Hurlbert and Beggs 1996; Latkin et al. 1995; Fischer and Shavit 1995)

The reason that there have been few studies of personal network structure on large numbers of alters is that respondents must report on the ties between alters. Getting respondents to assess all alter-pair combinations is a tedious task that increases geometrically as alters are added (see Figure 1). Even though the process of evaluating an alter pair tie is relatively quick (rarely more than five seconds), it can easily take an hour for a respondent to complete all 990 alter pair evaluations for a 45 alter network. Combining this with information elicited about the respondent themselves, and about each of their network alters, it is difficult for researchers to justify obtaining structural data on large

numbers of alters. On the other hand, given that elicitation of very few network alters typically result in mostly strong ties, the structural data from these samples are less interesting than large samples of alters that demonstrate structural variability. A method that maximizes structural variability while lowering respondent burden would advance the field of personal network analysis.

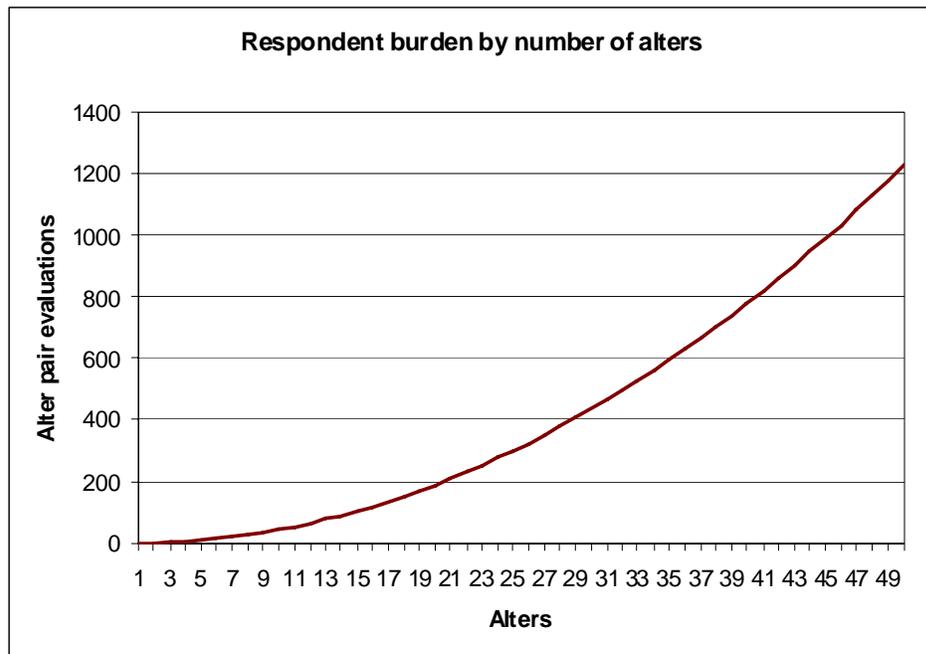


Figure 1. Respondent Burden by number of alters.

METHOD

The data for this study were generated as part of a grant to develop a web-based personal network intervention for adolescents at risk of smoking. This technology is founded on literature that identifies social influences as the primary factor explaining adolescents transitioning from non-smokers to experimenting, and experimenting to regular smokers (Flay et al. 1994). Ennett et al. (1994) demonstrated the importance of the network structure of peer groups on smoking, although the alters for that study were constrained to be from a whole network consisting of a set of schools. By visualizing the structure of their personal network and the structural placement of key alters, including smokers, adolescents can then use simulation tools to understand both the effect of smokers on them and the consequences of changing those relationships. This software is viewed as a potential interface for other intervention tools as well.

All participants were college freshman and sophomores, as we wanted respondents who were as close to high school age as possible. The study began with an EgoNet interview of 100 respondents, 50 smokers and 50 non-smokers. EgoNet is a personal network data collection and analysis package freely available through the Internet (<http://survey.bebr.ufl.edu/egonet/>). It consists of two programs, one for creating a study and one for running it. A study consists of four sections: questions asked of the respondent about themselves, questions asked to elicit a set number of alters, questions asked of the respondent about each alter, and a question about the tie between each unique pair of alters. The last module is used to generate adjacency matrices for structural analysis.

EgoNet uses the adjacency matrix to generate a network visualization based on the open source software library JUNG (Java Universal Network/Graph Framework), developed primarily at the University of

California at Irvine. EgoNet also calculates several structural measures (degree, closeness and betweenness point centrality, degree, closeness and betweenness network centralization, the number of components greater than size 2, the number of dyads, the number of isolates and the number of cliques). These measures can be viewed for an individual personal network in EgoNet or output as a summary file that combines respondent data, compositional data about all alters and structural data about the network in a comma delimited file with one line per respondent. All 100 respondents completed a 45 alter EgoNet study. The last module of EgoNet requires the respondent to evaluate all alter-alter pairs, in this case 990 ties. For this particular study, respondents typically finished in less than two hours. Respondents were paid \$30 to complete the EgoNet study and submit to a short interview following the study where they were asked questions about their personal network visualization.

EgoWeb was developed to reduce respondent burden and to deliver a personal network interview over the web. EgoWeb relies on dynamic network visualization, whereas other network visualization packages (including EgoNet) expect as input a completed adjacency matrix. EgoWeb uses a visual interface for collecting personal network data and redraws the network visualization with the addition of each alter. The purpose of the study was to test the EgoWeb interface that was designed to reduce respondent burden and to make the interface more appealing to respondents.

Unlike the EgoNet study that relied on a free-list of 45 alters, the EgoWeb study was designed to elicit alters in such a way as to maximize network structural features early on. Respondents were asked to list a single alter, but not one who is closest to them. They were then asked to name someone they knew who also talked to that alter. As these alters were added to EgoWeb a dot was placed on the screen with the alter's name below it and a line placed between the alters and the dot representing the respondent, indicating a network tie.

Next respondents were asked to name an alter who they knew, but who did not talk to any of the other two alters already depicted. They were then asked to name an alter who talked to the one just mentioned. If they couldn't think of someone that talked to that alter, that alter was an isolate. This process of naming pairs or singles that were unrelated continued until the respondent couldn't name any more. The idea was to force the respondent to nominate people from the variety of whole networks to which they belonged. Only when all such groups were exhausted did the respondent proceed to the next stage.

Once all whole networks had been represented, respondents were asked to name more alters until the visualization contained 45 names. At this time they were asked to concentrate on very close alters, that is, those they would not want to leave out. Close alters were avoided for the first part of the elicitation task as close alters tend to be bridges in a network and would make it difficult to name pairs of alters that were not tied to each other. For this part of the elicitation, as a new alter was named the respondent clicked on existing nodes, selecting those alters to which the new nominee was tied. The respondent indicated when they were finished making those ties and the visualization was refreshed and the respondent could list a new alter. This process continued until 45 alters had been named.

The EgoWeb elicitation differs from the EgoNet elicitation in several ways:

1. Respondents are forced to list pairs of unconnected, and less-close alters before listing more close alters. This maximizes network structural features.
2. Respondents see the visualization as they enter alters. This no doubt affects who they list next. Some respondents may actually use the visualization to try to fill out groups they see clustering on the screen.

3. Respondents only provide existing ties using the visual interface. In EgoNet all ties must be evaluated, including null ties. In a less cohesive network, null ties can easily represent the majority. This dramatically increases respondent burden in EgoNet.
4. Respondents in EgoWeb determine which ties they evaluate; often based on their perceptions of groupings they see through the visualization. For example, a respondent creating ties for a co-worker may easily avoid making any tie to family if they know there is no link between the work and family alters. Shifting control of which ties to evaluate from the researcher (via EgoNet) to the respondent (via EgoWeb) is perhaps the biggest change. With EgoWeb the researcher must rely on the respondent to provide the ties as the respondent burden is determined entirely by the respondent, rather than by the researcher.

Of the 100 respondents to the EgoNet study, 50 were selected to pilot the EgoWeb study and were paid \$60 to complete it via the web. The EgoWeb study was a much shorter version of the EgoNet study, including only the 45 alter elicitation using the dynamic visualization described above, and a question about alter smoking. The purpose was to get respondent feedback about the two methods and to test differences in selected structural measures between the two methods.

RESULTS

Given that the 50 respondents from this study used both the EgoNet and the EgoWeb interfaces, they were in an ideal position to make a comparison. It is not surprising that 70 percent of those respondents preferred EgoWeb. On average, the EgoWeb task consisting of simultaneous alter elicitation and alter-alter tie evaluation took about half the time of the EgoNet alter elicitation and alter-alter tie evaluation modules combined. The EgoWeb interface would have no effect on questions asked of the respondent about themselves or about their alters.

Specific comments about the comparison were more varied. Most respondents found it easier to point and click rather than to muddle through the arduous task of responding to the 990 alter-pair evaluations. Most also found the visualization interesting and appealing. One respondent suggested that having the visualization on the screen helped her to recall certain respondents. This is a factor that must be examined further. It is unclear whether the feedback is a positive influence, helping the respondent describe their network the way they perceive it to be, or a negative by stimulating them to fill out clusters that would otherwise be less represented.

On the negative side, some respondents actually felt it was harder to think of people to list, having the natural flow of the free-list disrupted by seeing their network structure. Many respondents found it difficult to read the names in tightly knit groupings of alters, even though the visualization allowed the nodes to be dragged aside. This is a technical issue that should be possible to correct by adding the capability to isolate an area of the visualization for expansion across the screen. Most respondents found it difficult to think of people they knew who did not talk to each other. This was to be expected given that it was designed to exhaust all of these groups.

Personal network composition refers to the summary characteristics of the alters who the respondent lists. This is in contrast to personal network structure that refers to the summary measures that capture the pattern of relations between those alters. While we cannot compare summary alter attributes of the two methods, we can determine to what extent respondents listed the same people in the two studies. Of the fifty respondents to EgoWeb, respondents on average used 22.4 (SD 6.5) of the alters they used in EgoNet. This represents only half of the alters from the free-list. The minimum that were the same was 1 and the maximum was 36. It is apparent that the EgoWeb interface generated a

much different set of alters. The differences could have been due to the elicitation method, the feedback from the visual interface or both. It is important to note that the two studies were conducted within a time gap of 2 months, so it is doubtful that the differences between the networks were due to actual changes in the respondent's personal network.

One obvious question is whether those that were substituted between the EgoNet and EgoWeb studies tended to be strong or weak ties. By comparing the average closeness score for each alter on a scale of 1 to 5, we found the average for those included in the EgoWeb study was 3.3 compared to 2.5 for those who were left out. The difference between these averages was significant ($p < .001$). As expected, both methods pick up the core network members and vary in the weak ties that are used to list a 45 alter network.

The structural differences between the two elicitation methods are summarized in Table 1. We expected that the EgoWeb method would elicit members of all the groups a respondent belongs to, which would in turn maximize both mean betweenness point centrality and mean betweenness centralization. In fact, these were the only two structural measures that were not significantly different. On average, these measures differed less than either degree or closeness centrality. While EgoWeb does result in higher numbers for both betweenness measures, the differences between these two measures is so highly variable (as demonstrated by the coefficient of variation) that significant differences between them were not found. There were, however, about 30 percent more network components of size three or more using EgoWeb than EgoNet. Although the bridging capability of alters was not significantly different, EgoWeb did result in more subgroups.

Table 1. Comparison of structural measures between EgoNet and EgoWeb elicitation for 50 respondents.

Variable	Mean – Egonet	Mean – Egoweb	Mean Difference	Coefficient of Variation of Mean Difference	Probability Difference > 0
Mean point degree centrality	10.90	7.50	3.41	1.00	0.0001
Mean degree centralization	36.20	31.20	4.99	3.25	0.03
Mean point closeness centrality	30.50	21.20	9.32	2.28	0.003
Mean closeness centralization	20.20	12.50	7.69	4.07	0.09
Mean point betweenness centrality	19.60	21.60	-1.97	-7.28	0.34
Mean betweenness centralization	22.40	23.80	-1.43	-15.11	0.64
Number of components	1.50	1.90	-0.44	-3.37	0.04
Number of isolates	1.50	0.70	0.76	2.95	0.02
Number of Cliques	46.90	32.20	14.62	2.53	0.008

All of the other measures exhibited significant differences. The most significant difference was mean point degree centrality, which was much lower in EgoWeb than in EgoNet. There were also about half as many isolates using EgoNet versus EgoWeb. The difference in degree centrality is likely due to the

fact that respondents to EgoWeb choose which alter-alter pairs to enter, rather than having to individually evaluate each one as they do in EgoNet. The lower number of isolates was probably affected by the respondents' use of the network visualization in EgoWeb as a cue once they got to the stage where they were entering individual alters. Respondents may have been less likely to enter isolates than members of groups they saw before them in the visualization.

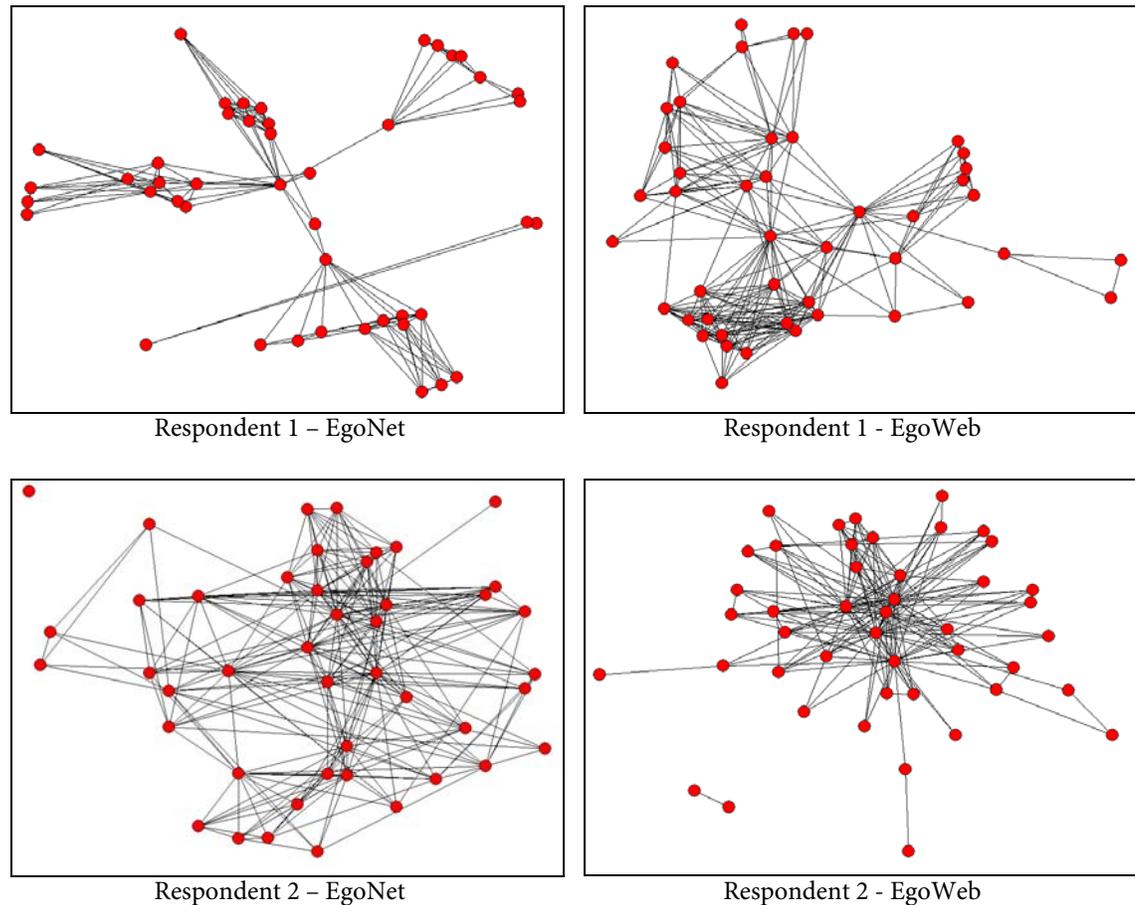


Figure 2. Selected network visualizations using EgoNet and EgoWeb.

Figure 2 compares personal network visualizations of two respondents using EgoNet and EgoWeb. Respondent 1 reflects most of the differences detailed in Table 1. In contrast, Respondent 2 reflects very few of the differences in Table 1. These two visualizations demonstrate the variability that one can expect using the modified elicitation of EgoWeb. Although most of the groups Respondent 1 listed in EgoNet are also represented in EgoWeb, the structural features appear quite different.

DISCUSSION

In this article we have compared two methods for collecting structural data on personal networks. The practice of visualizing personal networks, that is unconstrained egocentric networks, is not common, and the analysis of structural data on personal networks is relatively unexplored. We believe that the variability in structural features of personal networks is a good source for explaining variability in many outcome variables. We also believe that future network applications will use visualizations of

personal networks as an interface, and thus understanding the best way to elicit these data is essential. Much research is needed in extracting valid and reliable data from respondents.

The two methods reported here, EgoNet and EgoWeb, are operationalized in two software programs. EgoNet is designed for researchers to collect personal network data across many respondents. It does not use a visual interface and collects structural data from respondents by presenting all possible alter pairs. EgoNet presents a network visualization to the respondent after all data are collected.

In contrast, EgoWeb is oriented toward an individual respondent. It is an attempt to create a personal network interface that can be used over the Internet to deliver network interventions. Given this constraint, it is designed to lower respondent burden and be visually appealing. This is done through the personal network dynamic visualization.

There are two big differences between the two methods. EgoNet relies on a set of elicitation questions typical of personal network research. EgoWeb is an attempt to extract the groupings a respondent belongs to without the bias of pre-conceived groupings used by researchers (e.g. family, work, church). McCarty (2002) found that groupings derived from alter-alter interaction often did not fall into these pre-conceived groupings. Thus, a method that forces respondents to list groups without using these prompts may better reflect their social environment without imposing a set of common cognitive groups upon them.

Perhaps the biggest difference between the two methods is that in EgoNet respondents are forced to evaluate each alter pair tie (990 evaluations for a 45 alter personal network) whereas the visual interface of EgoWeb allows the respondent to use the visualization to tie a new alter to those already depicted. Although this is designed to reduce respondent burden by allowing respondents to avoid evaluating null ties, it makes it possible for unmotivated respondents to leave out ties that may not be null. The consequences of that are evident in that EgoNet generated 493 ties and EgoWeb 340 ties out of the possible 990. EgoWeb results in significantly ($p < .001$) fewer ties.

Although the visual interface of EgoWeb reduces respondent burden by about half, it results in a much different structure. It is unknown which of the two structures is closest to the existing communication between the alters, yet it is likely that the EgoNet structure is more reliable given that each alter tie evaluation must be made. For research, particularly in cases where respondents are compensated, it is still advisable that respondents evaluate all alter pair combinations. This ensures that the respondent burden is the same for all respondents.

It is unclear, however, that the modified elicitation that attempts to elicit all groups a respondent belongs to before listing the remaining alters is advisable. In hindsight, it would have been useful to have the 50 respondents who had used EgoWeb evaluate all 990 alter pair combinations and see how close that structure is to EgoNet.

While the dynamic visualization presents some problems, for the purposes of a network intervention, some form of this interface will likely be necessary. More research must be done to reduce respondent burden and to maximize the validity and reliability of the structural data that result. For example, several algorithms may be used to predict ties based on existing data. One approach would be to have the software assume ties by completing triads. Another would be to use attributes of alters, such as the relationship category, to assume ties. Any algorithm that assumed ties would have to test some sample of them, introducing more ties to evaluate if the assumptions were proven incorrect.

The future application of personal networks as a tool for intervention will undoubtedly involve some type of visual interface as has been tested here. We assume that a respondent that is motivated, that

is one who hopes to get something from the intervention, would take the time to ensure that the structure of their network is as accurate as possible, and that a user of a personal network intervention would be motivated. Further research will hopefully yield an interface that is both entertaining and accurate.

REFERENCES

- Brewer DD. 1997. No associative biases in the first name cued recall procedure for eliciting personal networks, *Social Networks*, 19 (4): 345-353.
- Brewer DD, Webster CM. 1999. Forgetting of friends and its effects on measuring friendship networks, *Social Networks*, 21(4): 361-373.
- Brewer DD. 2000. Forgetting in the recall-based elicitation of personal and social networks, *Social Networks*, 22(1): 29-43.
- Brewer DD, Garrett SB . 2001. Evaluation of interviewing techniques to enhance recall of sexual and drug injection partners, *Sexually Transmitted Diseases*, 28(11): 666-677.
- Ennett, Susan T.; Bauman, Karl E. 1994. The contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking, *Journal of Personality & Social Psychology*, 67(4): 653-663.
- Fischer CS and Shavit Y. 1995. National Differences In Network Density - Israel And The United-States, *Social Networks*, 17(2): 129-145.
- Flay Br, Hu Fb, Siddiqui O, Day Le, Hedeker D, Petraitis J, Richardson J, Sussman S. 1994. Differential Influence Of Parental Smoking And Friends Smoking On Adolescent Initiation And Escalation Of Smoking, *Journal Of Health And Social Behavior*, 35(3): 248-265
- Haines VA, Hurlbert JS and Beggs JJ. 1996. Exploring the determinants of support provision: Provider characteristics, personal networks, community contexts, and support following life events, *Journal Of Health And Social Behaviour*, 37(3): 252-264.
- Killworth, PD, Bernard, HR and McCarty, C. 1984. Measuring Patterns of Acquaintance, *Current Anthropology*, 25:381-397.
- Latkin CA, Mandell W, Knowlton AR, Vlahov D and Hawkins W. 1998. Personal network correlates and predictors of homelessness for injection drug users in Baltimore, Maryland , *Journal Of Social Distress And The Homeless*, 7(4): 263-278
- Latkin C, Mandell W, Vlahov D, Knowlton A, Oziemkowska M and Celentano D. 1995. Personal network characteristics as antecedents to needle-sharing and shooting gallery attendance, *Social Networks*, 17(3-4): 219-228.
- McCarty, C., Bernard, HR, Killworth, PD, Johnsen, E and Shelley, GA. 1997. Eliciting Representative Samples of Personal Networks, *Social Networks*, 19:303-323.
- McCarty, C, Killworth, PD, Bernard, HR, Johnsen, E and Shelley, GA. 2000. Comparing Two Methods for Estimating Network Size, *Human Organization*, 60:28-39.
- McCarty, C. 2002. Measuring Structure in Personal Networks. *Journal of Social Structure*, 3:1.