

# Conceptual and Empirical Arguments for Including or Excluding Ego from Structural Analyses of Personal Networks

**Christopher McCarty**

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

**Amber Wutich**

*Department of Anthropology and Bureau of Economic and Business Research,  
University of Florida, Gainesville, FL*

*The structural properties of personal networks are potentially fruitful variables for explaining differences in attitudes, behaviors and conditions across individuals. When researchers apply structural measures to personal network data, they must decide whether to include or exclude ego from the adjacency matrix. This research note discusses several conceptual and empirical issues that should be considered in making that decision.*

Most personal network research over the past forty years has focussed on network composition (summaries of alter attributes) rather than network structure (analysis of the pattern of ties between alters). While compositional analyses yield vital information about the network and how it impacts the respondent, structural properties of networks offer a unique perspective and are a worthwhile pursuit (McCarty, 2002). As researchers create new software that makes it easier to construct studies that collect alter-to-alter tie evaluations from respondents (the basis for structural analyses of personal networks), structural features will add to the set of network features that are used to explain respondent attitudes, conditions, and behaviours.

There is, of course, an established tradition of constructing egocentric networks within sociocentric networks. For example, Burt's notion of structural holes is a concept derived from looking at the

egocentric network of individuals within a corporate setting (Burt, 1992). While this approach is useful, it is vastly different than the case of personal networks where the list of alters is constrained only by the existence of a link to ego and may span across many groups. We suggest that personal network research should be used primarily to determine the effects of ego's network on ego, or to compare differences in interaction patterns across egos.

Given the novelty of the structural approach in personal network studies, we expect questions to emerge over how personal network data and structural measures should be handled. For instance, when analysing personal network structure, particularly for personal networks with more than 30 alters, researchers will have to determine whether or not ego's ties to her alters should be included in the adjacency matrix that is the input for structural analyses. In this paper, we will explore the conceptual and empirical issues that go into making this decision.

### **CONCEPTUAL ISSUES**

Analyses of whole (sociocentric) networks, though often complex, are conceptually straightforward. A whole network is a group of actors who, to the outside observer, appear more likely to interact than a randomly selected group of actors of the same size. For example, we expect that the 25 members of a drama club in a high school will have more opportunity and reason to interact than 25 randomly selected students from the entire high school population. Network structural measures applied to the drama club are likely to show patterns of interaction where the measures applied to the group of randomly selected students will not.

The conceptual issues surrounding personal networks are more difficult. Nobody considers the analysis of the composition of personal networks to be problematic. We understand how the proportion of a person's network that is female, White, or provides emotional support, or the average age and strength of tie of alters, might affect that person's attitudes, conditions, and behaviors. It is less clear whether the network structure derived from the assessed interactions between alters is meaningful. In other words, is the structure of a personal network of any practical importance?

The interaction within a social group where membership is defined by having some relationship to a single person is meaningless, unless the impact on or of that person (ego) is somehow involved. One may know 300 people, but those 300 people may span large geographic and socioeconomic distances.

Without reference to ego, they are conceptually somewhere in between the examples of the drama club and the random selection of students given above. They are more likely to interact than a random sample of people, but those interactions are hardly interesting or meaningful. This may seem an obvious point, but we believe that structural results should be interpreted only in terms of how the network affects ego or how ego affects the network. To interpret them as a network with innately meaningful structure, as in a sociocentric network, assumes that members perceive the personal network to be social group – an assumption not generally borne out by the data.

The approach researchers use to include ego in data collection and analysis should be determined by the needs of the research question. There are three ways to include ego in the analysis. First, one can leave ego out of the adjacency matrix. Second, one can include ego as a network member, forcing a tie from ego to all other members. Finally, ego can be included as a member of the network, using a tie definition that allows for null ties between ego and each alter. We will discuss conceptual issues involved in each of these in turn.

**Leaving ego out**

The first option is to exclude ego from the adjacency matrix. This is, in fact, the approach suggested by Scott (1997). At first glance, conducting analysis on a personal network without ego included may appear strange. It is, after all, ego's network and would not exist if not for ego. However, the social environment in which we live is, for many things, not brokered by us, even though we are at its center. One obvious example is gossip. Gossip tends to transfer within the network of people we know without our control over that flow. It can affect us without our knowledge, facilitation, or control.

Another example is social support. Consider a person who is elderly and in need of daily home care. The structure of that person's informal support network may determine if such care is administered informally, by network members, or by strangers, through some formal home care organization. The support network may not be brokered by the elder concerned, even if he is at least partially responsible for its structure. In cases where ego is seeking support, she may not be driving the avenues of delivery of that support. For many things ego is a passive receiver of information and resources from the network. The structural pattern of the network without ego's influence provides a unique picture of the social material ego has to form their attitudes, conditions, and behaviours.

**Including ego, with forced ties**

Consider the second option, that is, to include ego in the network but force a tie to all alters. Any time the same prompt is used to elicit ego's network and to determine alter-alter ties, ego must be linked to each alter. In such a case, null values cannot exist between ego and the alters. For example, if the researcher asks ego to list every person she conducts business with, and then asks ego if each pair of alters conduct business together, then ego would necessarily be tied to every alter in the network. Similarly, when the prompt used to elicit alters is subsumed by or included in the definition of ties used to evaluate connections between them, ego will always be linked to each alter. For example, if the researcher asks ego to list every person she conducted business with in the last month, and then asked ego if each pair of alters had conducted business together in the last year, then ego would necessarily be tied to each alter.

Connecting ego to every alter is a very intuitive approach since we usually are interested in using the same kind of tie to elicit ego's network and to find out about interactions within that network. Given the compositional analysis approach, we are used to thinking of personal networks as stars. It seems natural that adding ties between alters should just result in an appended star structure.

By virtue of the fact that ego is now tied to all alters, ego will affect network structure more than any other alter – perhaps an unintended consequence. The focus is no longer on how the network affects ego, but instead on how ego affects the network. However, a personal network is not an independently existing social group whose structural patterns we want to predict.

A primary reason for collecting personal network data is to understand how the network impacts ego. By including ego, this impact has largely been removed because we can now only analyze issues where ego's overwhelming influence is a valid question. There are examples of this. If we imagine ego being the focal point of some special knowledge or condition, we might want to know the impact of ego on all of the alters we elicit, given the pattern of relations between those alters. For instance, if ego has a condition that is typically transmitted in every instance of face-to-face contact, such as a highly infectious disease, we can imagine that understanding the structure of face-to-face interactions in ego's personal network, and how structures differ across egos, might be of great interest to some researchers.

### **Including ego, allowing null ties**

The third option is to include ego, but allow for null ties between ego and some subset of alters. This seems odd at first, but it may be the most appropriate approach to take to compare patterns of interaction across egos. By allowing the tie to be null, we separate the elicitation task from the tie definition. We can think of many examples where ego knows someone, but has no tie connection. For example, we may choose to do an analysis where we ask ego to list the 50 people he knows best, and then to assess the likelihood of each pair of alters discussing politics. In this case, there may be ties between alters where politics are discussed (e.g. father and son), but ego is unlikely to discuss politics with either of those alters. Again, the focus is now not on how this network impacts ego, but on how ego, as any other alter, impacts the network. The tie definition should reflect that.

We can imagine some cases where this approach may be quite useful. Again consider an epidemiological study, this time examining the personal networks of HIV positive IV drug users. In such a case, we may be interested in using instances of needle sharing or sexual relationships to define ties between alters. Although ego may know the people in his network, ego probably does not share needles or have sex with all of them. In this case, we can examine how the potential for the spread of HIV is mediated by structural variability in the personal network. We might be able to design interventions that take this network structure under consideration. Under this scenario, ego plays a vital role in the structure of the personal network and must be included, but allowing for null ties.

### **EMPIRICAL ISSUES**

Aside from the conceptual issues, the application of structural measures to personal networks presents empirical problems as well. First, we must keep in mind that we are asking respondents to report on the nature of the relationship between alters. We know that respondents can assess, in general terms, whether their network alters know each other. This is demonstrated by the sensible groupings of alters that respondents can identify using network visualizations and personal network adjacency matrices. If respondents could not make these assessments, the resulting structural patterns they see through network visualizations would appear arbitrary. Further, McCarty (2002) found respondents' alter-alter tie evaluations to be reliable when asked to reassess tie evaluations they had already made.

We are less confident in the ability of respondents to assess more subtle relations between alters. For example, respondents may be able to assess the strength of relationship between alters as non-existent, weak tie and strong tie, but may not be able to assess those ties on a five or ten point scale. Since the data quality is of such crucial importance in network studies, where even a few missing ties could significantly alter network structure, we recommend proceeding cautiously with such studies. The ability of respondents to make these evaluations would vary depending on the size of the personal network and their familiarity with their alters. If the elicitation were limited to close family, the respondent may be able to provide more detailed information. However, it is not enough for respondents to report accurately on some alter-alter ties. They must be able to report accurately on all of them. The level of knowledge required to make that assessment must be driven by the tie the respondent knows the least about.

We are cautiously optimistic about the prospect of personal network researchers finding ways to study subtle or asymmetrical relationships with relatively small personal networks or networks where all of the alter-alter ties are extremely well-understood by ego. For instance, if we elicit ego's network of closest family members and how much money alter A has lent alter B in the last year, we may be able to ascertain a fairly accurate picture of the family loan network, using asymmetrical, interval-level data.

Because of these limitations on what respondents can reasonably report, we believe that the typical adjacency matrix for a personal network will be symmetric, not directed, and will have at most three levels of tie strength (no tie, weak tie, strong tie). This precludes the use of some structural measures. However, others are still valid. We will proceed with a discussion of how structural measures could be used to study adjacency matrices that exclude ego and adjacency matrices with ego included, forcing ties to all alters. In terms of execution, there is no difference between adjacency matrices that exclude ego and those that include ego but allow for null ties between ego and alters. Therefore, we will discuss the feasibility of using nine common network measures for adjacency matrices that do and do not force ties between ego and all alters.

**Density** is the proportion of existing ties out of all possible ties. It is valid for all approaches. Adding ego and forcing a tie increases the number of ties by the number of alters over the no ego approach.

**Degree Centrality** for a given alter is the number of alters they are directly connected to. It is valid for all approaches. Adding ego and forcing a tie increases the point centrality of each alter by one compared to the no ego approach.

**Closeness Centrality** for a given alter is the inverse of the distance from that alter to all other alters. Personal networks can (and often do) have network isolates. Closeness centrality is not meaningful with unconnected graphs (the presence of isolates or components). Ego must be included for closeness to be reliably calculated. While we may be able to calculate closeness centrality for some respondents, when we are comparing across respondents we cannot count on a valid closeness centrality result.

**Betweenness Centrality** for a given alter is the number of geodesics (shortest paths) between all alters that the alter is on. Although this is a valid measure when ego is included and a tie is forced, it becomes strongly correlated with degree centrality. In a graph without ego, there is opportunity for alters to serve key bridging roles. When ego is included they, by default, lie on the most geodesics, except when alters have direct ties.

**Components** are connected graphs within a network. When ego is included and a tie is forced, the graph is by definition connected and there can only be one component.

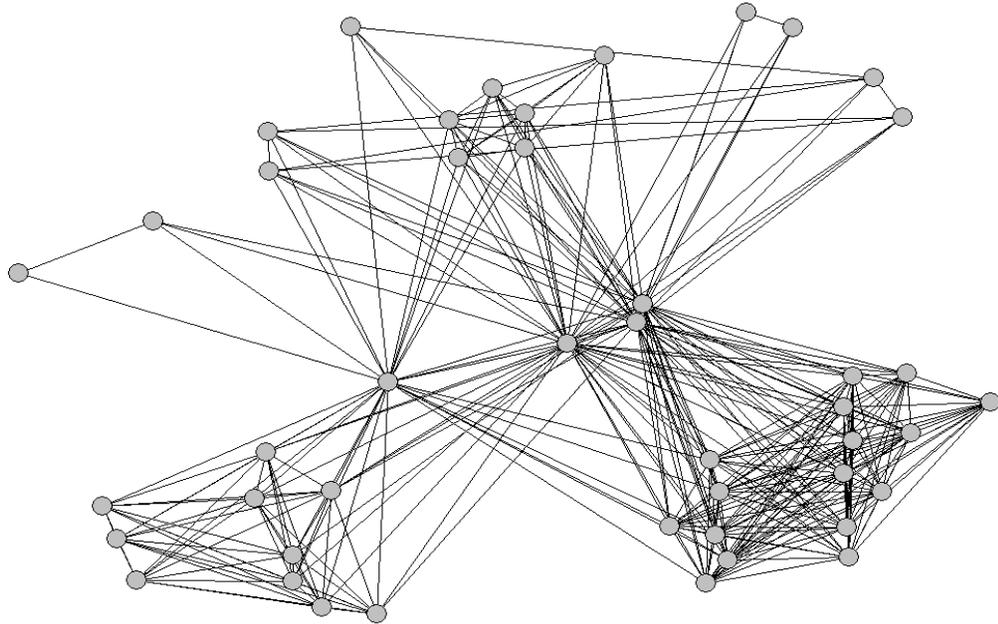
**Cliques** are maximally complete subgraphs. With ego networks, the difference between one clique and another is often the substitution of a single alter. Given that ego is automatically a member of every clique, the addition of ego and forcing a tie does not generally change the number of cliques. The same holds true for other measures in the same family (n-clique, n-clan, k-plex).

**Core-Periphery** attempts to identify a network core of alters who are all mostly connected. The procedure is less stringent in its definition of a subgroup than the clique routines and it results in only one core. Unlike the other procedures, the addition of ego may serve to bind groups that otherwise would not. Conceptually, this procedure may make more sense with ego included and forcing a tie to all alters as it defines the core group of people on which ego relies.

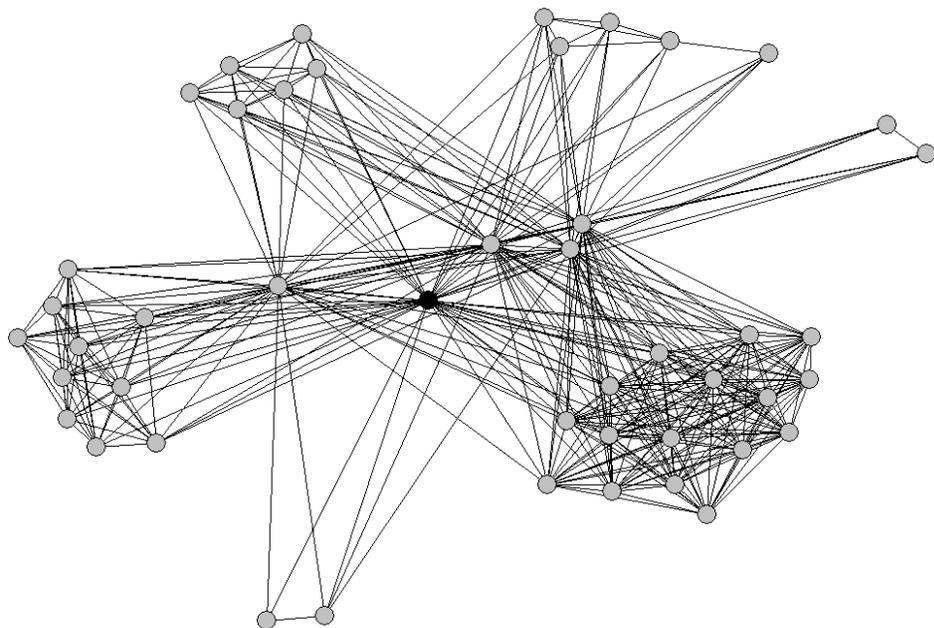
**Structural Equivalence** clusters alters together based on the structural role they serve in the network. Adding ego and forcing a tie adds an alter with what is usually a unique position, that is, they are connected to all alters. Adding ego does not change the structural role of the other alters.

**Network Visualization** routines are used to graphically depict the relations of nodes to each other. Although there are many algorithms for displaying network data, most will be affected the same way

by the addition of ego and forcing a tie to all alters. Ego will be at the center of all connections. Like core-periphery, this may be a helpful anchoring for respondents who are being interviewed about their network. Figures 1 and 2 show a network of 45 alters collected in a recent study. It is visualized with Netdraw, the first excluding ego and the second including ego. The structures look very similar.



**Figure 1:** Network Visualization Excluding Ego



**Figure 2:** Network Visualization Including Ego and Forcing a Tie to All Alters

**CONCLUSION**

These conceptual and empirical issues can be summarized as follows. If we are interested in the impact of social networks on ego, then analyses should be limited to adjacency matrices that do not include ego. We should think of ego as a passive receiver of information and resources that are transmitted across the network. In this case closeness centrality, and any other measure that requires a connected graph, cannot be calculated.

If we are interested in the impact ego has in brokering their network, then we should include ego. We should think of ego as an active participant in information and resource exchange. Many of the structural measures will be functionally the same as the case where ego is excluded. If a tie is forced, components will be meaningless, and betweenness centrality will reduce to degree centrality.

Finally, we may be interested in how ego impacts their network, but using a definition that allows for null ties. The empirical issues are the same as those for adjacency matrices without ego, that is, closeness centrality cannot be calculated.

**REFERENCES**

- Burt, R.S. 1992. *Structural Holes*. Cambridge: Harvard University Press.
- McCarty, C. 2002. Structure in Personal Networks. *Journal of Social Structure* 3:1.
- Scott, J. 1991. *Social Network Analysis: A Handbook*. London: Sage.