Pre-print of an article scheduled to appear in 2013 Social Networks

How many names are enough? Identifying network effects with the least set of listed contacts

March, 2013

Jennifer Merluzzi
Tulane University

Ronald S. Burt
University of Chicago

Abstract
How many names are enough to reveal network effects using a name generator for network analysis? We analyze network data from two large organizations varying in complexity. We ask how much the network association with achievement is strengthened by adding another name to the recorded list of each person's sociometric citations. We conclude that five names is the cost effective number of sociometric citations to record. The network association with achievement weakens quickly with fewer names, especially in a more clustered network.

Keywords: Name generators; Personal networks; Survey methods; Returns to brokerage

Introduction

This note addresses a practical question inherent to designing a name generator for a network analysis: How many names should be solicited? If studying advice networks, how many advisors? If studying collaboration networks, how many collaborators? There is methodological work offering guidance on what kind of relations to ask about (e.g. frequency versus emotional closeness, obligatory versus discretionary, see Marsden and Campbell, 1984; Baron and Podolny, 1997; Burt, 1997, 2010: Appendix A; Bearman and Parigi, 2004; Marin and Hampton, 2007), but almost no work answers the "how many names" question. Burt (1986) argues for more than three names in the General Social Survey (GSS) because co-workers tend to be named fourth or fifth after close family have been named (cf., Wellman, 1979: 2010), but we are particularly interested in network analyses inside organizations. Such analyses differ in two broad ways from general population surveys like the GSS. First, name generators often have an upper limit greater than the 5-person limit in the GSS. Second, rich data available on work achievement closely related to work relationships can be used to sort through individuals for networks associated with high versus low network advantage.

In this paper, we analyze network data on two large organizations varying in complexity, one in which everyone can reach one another within a few choice links versus one in which the population is balkanized into regional clusters. We ask how much the network association with achievement is strengthened by adding another name to the list of sociometric citations recorded from each respondent. Knowing that people with many non-redundant contacts have a network advantage (Burt, 1992; Burt, Kilduff, and Tasselli, 2013), we analyze how contact redundancy varies as more sociometric citations are recorded from a survey respondent. We conclude that five names is the cost effective number of sociometric choices to record.

Research Background and Motivation

In the last two decades, an abundance of evidence has emerged on network advantage in organizations (see Brass et al., 2004; Burt, 2005; Burt, Kilduff, and Tasselli, 2013). As these results have become increasingly well documented, network researchers have sought improved methods and data collection strategies. One avenue of methodological interest has focused on reducing respondent burden in conducting network studies (Golinelli et al., 2010; McCarty, et al., 2007; McCarty et al., 1997). Respondent burden represents the time and energy individuals must dedicate to complete a network survey. Multiple methods – from telephone surveys to traditional pen and paper – have been utilized to collect data on an individual's personal network often relying on a series of name generator questions (see Laumann, 1966 for early example of name generators, but see also Marsden, 1990, 2005; Bidart and Charbonneau, 2011; Marin and Hampton, 2007 for more recent details on the method). Asking individuals to list a series of names of others (alters) across a variety of contexts, such as whom they discuss work with, go out to dinner with, would ask to borrow money from, discuss important matters with, provide social network researchers with a view of relationships among individuals in a target population. These targeted populations can be clearly bounded (whole network) or represent a subset highlighting personal (ego) networks defined by the name generators.

Completing a network survey requires respondents to list the names of people he or she interacts with typically across a variety of contexts (e.g., social, support, exchange, kin, supervisory) and then also provide details on the relationships among their alters as well as characteristics identifying the nature and strength of these relationships. To do so, necessitates time and motivation from the respondent. As Marin & Wellman (2010: p. 15) comment: "These surveys or interviews can be difficult and burdensome for both respondents and researchers. Ego network surveys especially – with their repetitive questions about each alter – can be long and boring. In addition, providing the information requested by researchers is difficult…" Continually, network researchers struggle with designing a survey that maximizes the cost-effectiveness of collecting network data by taxing the respondents with the least burden without jeopardizing the response rate and data quality provided by those respondents.

Toward this end, researchers have recently advocated approaches to reduce respondent burden with minimal impact on data quality and interpretation. The results have been varied with some key problems emerging. First, individuals do not consistently interpret questions in the same way. Relying on single name generators to save time becomes problematic because individuals may interpret a particular name generator question any number of ways. For instance, taking one of the most commonly used single name generator questions of “discussion of important matters” (Burt, 1984 from the General Social Survey), Bearman & Parigi (2004) revealed a variety of respondent interpretations and demonstrated a strong topic-alter dependency, where individuals define important matters differently depending on their relationship to a particular alter. Re-examining the results reported by McPherson et al. (2006) that found American’s having smaller discussion networks than in the past, Hampton et al. (2009) discovered possible changes over time in the way that individuals interpret the word “discuss” thus potentially biasing responses. And, in their study of four Boston residential communities, Marin and Hampton (2007) found two (out of six) name generators (“discussion of important matters” and “friends”) able to reliably serve as a single name generator for their study, but not consistently across all the network measures that they evaluated. Further, none of the six name generators they asked could consistently predict different role relationships, such as exchange or support.

In addition to respondent interpretation problems, it’s not clear that the data provided when the survey is abbreviated is in fact representative or does not lead to other researcher-related interpretation problems. For example, Kossinets (2006) cautioned about the risk of non-random missing data in the use of fixed choice methods, where the maximum number of names allowed in answering each name generator question is constrained (compared to free listing, where respondents can name as many alters as they see fit). A problem they describe is that individuals who would arguably have a greater likelihood of nomination by contacts in a network could be artificially inflated under fixed choice constraints. Additionally, Marin (2004) showed that when asked to list contacts, respondents first name those individuals with whom they share stronger relationships, although she also found that neither the multiplexity (variety in the types of relationships) nor the duration of the relationships mattered in terms of respondent recall. Along these same lines, respondent recall problems and forgetfulness (Bernard et al., 1984; Brewer, 2000; Brewer and Webster, 1999; Bell, et al., 2007; Feld and Carter, 2002) can plague network data, making it difficult to abbreviate the collection method for fear of gathering an overly biased sample.

Nonetheless, studies continue to put forth viable short cuts to consider, particularly when handling a research question that is specific. For instance, McCarty et al. (2007) tested four different short cut methods including dropping alters at the end of a free list, randomly dropping alters, randomly dropping links, and using transitivity to predict ties. They found instances where they could sufficiently capture the whole network structure with less than 20 alters and sometimes as few as 10 alters with the use of randomly dropping alters. In the study of four residential communities in Boston, Marin and Hampton (2007) tested both using a reduced selection of two of the most reliable of the six name generators (MMG approach) and using a random sampling of alters from the six name generators (MGRI approach) and then compared both of these methods to the data produced by the full six name generator design. Based on this, they found examples of a single name generator around friendship or support reasonably approximating the results found in the multiple name-generator approach, particularly when measuring network density and size. Borgatti, et al. (2006) introduced four types of random error into a network analysis including node and tie addition and deletion to understand the effect on four different measures of centrality. They found no significant difference among the four different centrality measures after introducing the errors, with node manipulations having less of an impact than tie additions or deletions. Similarly, Costenbader and Valente (2003) showed stability in various centrality measures even when half of the data was randomly missing.

In addition, others have explored alternative time and cost reducing collection methods as another way to reduce respondent burden. Kogovsek et al. (2002) compared data collection using automated telephone methodology compared to more time consuming face to face interview techniques to see if collection time and expense could be reduced without significantly impacting data quality. They found that the less expensive, less respondent time intensive telephone collection in fact produced more reliable results which they attributed to the anonymity of the interaction that allowed individuals to be more open in their responses and answer on their own time rather than a pre-determined interview arrangement. Schiffer and Hauck (2010) created and tested a new methodology of collecting network relationships through visual exercises done in small groups (Net-map) demonstrating reasonable reliability from small group settings in lieu of network surveys.

Finally, issues of informant accuracy in recall have long been raised in conducting network analysis (see Bernard, Killworth, Kronenfeld & Sailer, 1984 for a summary of a series of experiments on this), raising questions of outer limit setting in collecting data. Brewer & Webster (1999) focused on the specific problems of respondent recall and forgetfulness and found that the recall data was reasonably sufficient in terms of approximating network density. Personal network density based on "recall only" data had a 0.92 correlation with the combined "recall and recognition" data, where recognition data came from a subsequent prompting of names based on a list of names provided to respondents after the free recall exercise. Similar levels of correlation applied to other measures such as reciprocity, transitivity, number of cliques, and four measures of centralization. They also found that while the general belief is that individuals forget weak versus strong ties more readily, the evidence on this is in fact mixed.

In sum, while caution clearly needs to be paid when taking short cuts in network research design, abbreviated network surveys can provide reliable results in many situations and across popular network measures like centrality and density of a network.

It is in this same vein that we situate our investigation. Our study focuses on the specific methodological question of how many names to elicit in a name generator survey. To date, determining the appropriate number of names to ask on a survey has emerged mostly through researchers defaulting to prior studies. Early studies typically asked three names, such as the local and regional marketing surveys conducted through Columbia University's Bureau of Applied Social Research (e.g., Lazarsfeld, et al., 1944; Katz and Lazarsfeld, 1955; Coleman, et al., 1957). The GSS recorded up to five names in interviews with its national probability sample of respondents (Burt, 1984; McPherson, et al., 2006). Kogovsek, Mrzel and Hlebec, 2010) capped it at two and found differences in network composition and structure to emerge when the network size was small or, depending on the type of relationship (emotional support seemed more sensitive). Further, while the merits of fixing the choice set versus allowing free listing has been debated (Feld and Scott, 2002; Kossinets, 2006; Marsden, 1990), even when presented with an unconstrained choice set, respondents almost unilaterally cap themselves at six contacts (Fischer, 1982; Burt, 1984; Marsden, 1990). It is generally unclear then how many names are sufficient to elicit on a network survey. If the researcher asks for too many, she eats up too much respondent time and risks irritating respondents before they complete the survey. But, if she asks too few she may inaccurately portray a network filled with cliques - especially if sampling on the first few names yields a biased result of an individual's closest colleagues that are typically listed first (Brewer, 1995; Marin, 2004).

Research Design and Data

We use construct validity to address the "how many names" question. Given widespread interest in the association between achievement and networks of non-redundant contacts, we ask how much more evidence of association is provided by recording additional sociometric citations from respondents. As each additional contact cited by a respondent is added to his network, and standard measures of contact redundancy are computed, what happens to the estimated network-performance association?

We use network data on two organizations with very different structures. Figure 1 contains sociograms of the two organizations. The organization to the left is balkanized into regional clusters. The organization to the right has a single-cluster, center-periphery structure. The structural holes between clusters in the first organization provide more opportunities for people to have non-redundant contacts. We want to know whether the greater opportunity matters for our answer to the "how many names" question.

The organization to the left in Figure 1 depicts individuals involved in sales, support, and regional operational functions of a new product launch in the Asia-Pacific operations of a large American software company. The Asia-Pacific organization consisted of three sales groups, a product support group, regional operations and supported by a central leadership team and administrative functions (such as finance and human resources) located across offices in the Asian region. Network data were obtained with a generic network survey instrument of name generators and name interpreters (Marsden, 1990, 2005). Respondents were scattered all over Asia so the instrument was administered via a webpage to which people were directed by an email message from their senior executive. Name generators asked for the respondent's supervisor ("Who is your immediate supervisor (person most responsible for your annual review and initial salary-promotion recommendation?") and people with whom the respondent had the most frequent and substantive work discussion ("More generally, who are the seven or eight people with whom you have had the most frequent and substantive work contact over the last six months?"). At the end of the survey, respondents were asked to complete an adjacency matrix describing connections between their contacts. Network datum z_{ji} is binary and symmetric, "1" if i cited j or j cited i or a respondent cited i and j saying that the two contacts were especially close to each other, and "0" otherwise. For more details on the organization and network survey, see xxx (2010, Ch. 3) or yyy (2013). The 162 respondents snowballed into 331 managers and senior employees. The average person was experienced, with over 5 years tenure and just under 39 years of age. Based on company records, 77% of the people held managerial rank, with a quarter of those considered senior executives.

[Insert Figure 1 and Table 1 about here.]

A similar network survey was administered in the software organization (to the right in Figure 1). The 113 survey respondents snowballed into a broader 507-person organization. Respondents worked in one geographic location, facilitating the one-cluster structure displayed in Figure 1. The population ranged across corporate functions and varied in age from their early thirties to mid-sixties. People were again asked to name their immediate supervisor. In addition, respondents were asked to name "the seven or eight people with whom you have had the most frequent and substantive work content over the last six months. Limit yourself to people with whom you have had direct contact. This can include email contact, but not email lists." In this organization too, the network data are binary and symmetric.

Given the similar instruments used to gather network data, it is not surprising to see similar response patterns in the two organizations. Table 1 shows that the modal response was eight cited contacts in both organizations with an average of 7.17 cited contacts in the balkanized Asia-Pacific firm and 6.95 in the one-cluster software firm.

Summary Result – Increasing Evidence of Network Advantage

Our construct-validity network-performance association occurs in both of our study organizations. Controlling for job rank, age, tenure, gender, divisions and geographic location, compensation in both organizations increases with the number of non-redundant work discussion contacts (tables not shown, but see xxx, 2010: Ch. 3; yyy, 2013, for details on the network-performance association in the two organizations).

The published evidence of network-performance association in our organizations is based on network measures computed from all cited work contacts. How much of the evidence would we find if the original field work had only recorded a few of the cited contacts? At what point in the sociometric citation process does redundancy emerge among the cited names? We see either of two response patterns likely: (1) Respondents first list their closest contacts in response to a name generator, then move outward to non-redundant contacts (Burt, 1984; Marsden, 1987; Brewer 2000; Marin, 2004). If this were the case in our two management populations, we should observe a weak association between achievement and non-redundant contacts across initial citations that strengthens toward the end of the sociometric order as non-redundant contacts are cited. (2) A second response pattern is that people begin by defining the range of their network with their non-redundant contacts, then add friends of those contacts (something like the reverse small world phenomenon, Killworth and Bernard, 1978). If this second response pattern characterizes our two management populations, we should observe a positive relationship between achievement and non-redundant contacts that strengthens with initial citations, then tapers off as respondents begin citing contacts redundant with their initial set of cited contacts.

[Insert Figure 2 about here.]

We find that the second response pattern characterizes our organizations. The summary evidence is in Figure 2. On the horizontal axis, contacts are ordered by the sequence in which they were named for work discussion. A contact was named first, then another was named, then a third was named, and so on. The horizontal axis begins with the second contact named because we need two or more contacts to compute the network measures.

The vertical axis in Figure 2 indicates the correlation between compensation and the number of non-redundant contacts as additional citations are added. We use Burt's

(1992: 51-53) “effective size” measure of redundancy, which measures the extent to which a contact is disconnected from the others in a person’s network. So, the result above “2” on the horizontal axis in Figure 2 is the correlation between compensation and networks composed of the first two work contacts each person cited. The result above “3” is the correlation between compensation and networks composed of the first three work contacts cited, and so on. If a person only cited two contacts, that person only has two contacts in the result over “3” (we are interested in how the evidence of a network-performance association improves as respondents are given the opportunity to cite additional contacts, not whether respondents make full use of that opportunity). The table below the graph in Figure 2 details the correlations in the separate organizations and after holding constant job rank.

We see two results in Figure 2. First, in both organizations, the association strengthens between non-redundant contacts and performance as respondents are allowed to cite more work contacts. More names yield stronger results. This result is consistent with the intuition behind eliciting more names on a survey. Second, the association increase is non-linear. The increasing correlation between non-redundant contacts and performance slows as more contacts are included. Later-cited contacts tend to be redundant with the contacts cited earlier. The table at the bottom of Figure 2 shows, at the second contact listed, that the network-performance association is negligible (0.01 correlation in the balkanized organization, 0.06 correlation in the integrated organization, 0.02 correlation across both organizations). When respondents can cite up to four work contacts, the network-performance correlation increases to 0.21, and the association is statistically significant holding job rank constant (2.7 t-test). Five recorded contacts raise the correlation to 0.28 and the t-test to a healthy 3.4. Six recorded contacts raise the correlation only slightly to 0.29 and the t-test to 3.7. The network-performance association increases a little more with seven and eight contacts recorded. The pattern holds for both organizations (balkanized or integrated network structure). In sum, the evidence of network-performance association with non-redundant contacts strengthens as additional sociometric citations are recorded, but less with contacts added after the initial handful.

More Detail on the Redundancy of Later-Cited Contacts

Figure 3 provides detail on the increasing redundancy of additional sociometric choices. The horizontal axis is again in sociometric order; first contact named, second, and so on through eight. The vertical axis is the contact-specific redundancy component aggregated across contacts in Burt's "effective size" measure of redundancy. Each added contact is evaluated for its redundancy with contacts already cited by a respondent. A new contact is zero redundant if he or she has no connection with any of the contacts the respondent already cited. A new contact has some higher proportion redundancy depending on the extent to which he or she is connected with contacts the respondent already cited.

[Insert Figure 3 about here.]

The bold lines through black dots in Figure 3 show how average redundancy increased across sociometric citations. The first contact cited is zero redundant by definition; there are no prior contacts. The second contact cited is a little more than 0.4 redundant with the first contact, the third is between 0.5 and 0.6 redundant with the first two contacts, and so on. Redundancy stabilizes after the fifth contact cited.

Redundancy is higher in the organization balkanized into regional clusters (graph to the left in Figure 3), but individual differences are greater within than between the two organizations. The average difference between the organizations is illustrated by the slightly higher bold line for the balkanized organization (0.7 redundancy in contacts cited after the fifth versus 0.6 in the contacts cited after the fifth in the integrated organization). Individual differences within each organization are illustrated by the spread between the solid and dashed lines above and below the bold, dotted line. Contacts cited by respondents in high-density networks — the third of respondent networks with the highest density — are more redundant than the average in either organization (compare solid lines to the bold, dotted lines). Contacts cited by respondents with low-density networks — the third of respondent networks with the lowest density — are less redundant than the average in either organization (compare dashed lines to the bold, dotted lines).

We dug a little deeper into contact redundancy by distinguishing citations to contacts within versus beyond the respondent's own department. The latter distinguish boundary-spanning ties (emphasized as gatekeepers in the early research on informal coordination within a firm, or, as network brokers in contemporary network theory – see, Allen and Cohen, 1969; Aldrich and Herker, 1977; Tushman, 1977; Brass, 1984, with a key transition to the more general network concepts; and Tortoriello and Krackhart, 2010 with a update on the early research). If a respondent limits his citations to work contacts in his own department, there is a redundancy to his contacts whether or not the contacts talk directly with one another. Disconnected contacts within the department would be less redundant than connected contacts within the department, but the mere fact that a respondent draws all his contacts from the same department is its own kind of redundancy indicator.

But, what constitutes an organization boundary – and thus, a boundary-spanning tie - is very different across the two organizations. In the Asia-Pacific organization (the left sociogram in Figure 1), boundaries are defined by geography: the sociogram shows seven clusters — a central hub in Singapore with six regional clusters radiating out from the hub like spokes on a wheel to ASEAN, Australia, China, India, Korea, and the U.S. Functional distinctions also exist between individuals in the organization, with some working in sales, some in product support, and some in operations. But the organization clustering most apparent in Figure 1 is between geographic regions. The one apparent exception is the three sub-clusters to the southwest in the sociogram, but those three sub-clusters are themselves national clusters within the ASEAN regional operation. In total, 84% of the cited work contacts are in the same regional cluster as the respondent citing them.

In contrast, there are no obvious boundaries between clusters in the software organization (displayed to the right in Figure 1). There are five broad functional distinctions between individuals in the organization — Business Development, Engineering, Production, Project Management, and Other Administration (legal, human resources, finance, etc.) — but work discussion moves across these distinctions about as often as it moves within them. 44% of cited work contacts are within the same functional department as the respondent citing them.

[Insert Figure 4 about here.]

Figure 4 shows that the probability of a respondent citing boundary-spanning contacts is little affected by sociometric order. In the integrated organization to the right in Figure 4, there is no average trend across sociometric order. The lines for high- and low-density networks crisscross one another. (As in Figure 3, high- and low-density are the top and bottom third of respondent networks.) A logit regression equation predicting which contacts are in the same department as the respondent citing them shows no association with sociometric order (0.71 z-score) and a not-surprising high tendency for the respondent's immediate supervisor to be inside the respondent's department (6.57 z-score). There is a more obvious pattern to boundary-spanning contacts in the Asia-Pacific organization (to the left in Figure 4), but the primary characteristic is the lack of such ties. Almost all work discussion occurs within each region. The respondents in high-density networks are the most likely to cite within their regional cluster, but there is no variation in that tendency across sociometric order (-1.32 z-score for sociometric order). Similarly, respondents in low-density networks are the most likely to cite into other regions, but the tendency is low over all (18%) and does not vary with sociometric order (-0.79 z-score). For both categories of respondents, the negative test statistics show a tendency for within-region work contacts to be cited early in the sociometric order. When the tendency is estimated across all respondents, it is weak but nevertheless statistically significant (-2.67 z-score with correction for autocorrelation between contacts cited by the same respondent). In sum, contact redundancy increases with sociometric order regardless of whether the contact is in the respondent's own group or spans across a boundary into another group. We created graphs like Figure 3 but with contact redundancy plotted for contacts cited by a respondent in his own group versus contacts he cited in other groups. The two graph lines are indistinguishable, showing the same higher redundancy for internal and boundary-spanning connections cited later in the sociometric order.

Conclusions

Our summary conclusion is that five names is the cost effective number of sociometric choices to record. Recording additional names will provide further evidence of association between achievement and network variables measuring contact non-redundancy (Figure 2). However, the additional names can be expected to be 60-70% redundant with the first five (Figure 3), which makes the additional burden on respondents of asking for more names questionable, especially if respondents are also asked to describe the matrix of connections among their cited contacts. Interestingly, boundary-spanning connections have no association with sociometric order. Later-cited contacts are no more or less likely to be boundary-spanning than the first five contacts cited. In sum, later-cited contacts are more redundant. While debate has emerged around the merits of fixed choice survey designs (Feld and Scott, 2002; Kossinets, 2006; Marsden, 1990), these results demonstrate that fixing the number of names is a viable shortcut when the goal is to estimate the returns to network advantage.

A caution is that the recommended five names can be affected by the population network structure. Of course, one does not know the network structure of a firm or target study group for certain prior to collecting network data, although some underlying knowledge may indicate a general expectation, particularly if it is one where an extreme outcome is anticipated – such as the two firms compared here. Three or four names would be sufficient to estimate returns to network advantage in a highly integrated context such as the software organization (to the right in Figure 1), whereas five or six names would be necessary in a more segmented network such as the Asian-Pacific organization (to the left in Figure 1). Asking for a sixth or seventh name may be prudent in contexts where the researcher suspects network effects may be harder to disentangle from other effects.

Overall, this simple analysis demonstrates a way for researchers to think about how many names to elicit from respondents, particularly under tight time demands. Encouragingly, five names are sufficient in each of our different study organizations to reveal strong association between achievement and network advantage.

References

Aldrich, H. and Herker, D. 1977. Boundary spanning roles and organization structure. Academy of Management Review, 2, 217-230.

Allen, T.J. and Cohen, S. 1969. Information flow in R&D labs. Administrative Science Quarterly, 14, 12-19.

Bearman, P.S. and Parigi, P. 2004. Cloning headless frogs and other important matters: Conversation topics and network structure. Social Forces, 83: 535-57.

Bell, D.C., Belli-McQueen, B. and Haider, A. 2007. Partner naming and forgetting: Recall of network members. Social Networks, 29: 279-99.

Bernard, H. R., Killworth, P., Kronenfeld, D. and Sailer, L. 1984. The problem of informant accuracy: The validity of retrospective data. Annual Review of Anthropology, 13: 495-517.

Bidart, C. and Charbonneau, J. 2011. How to generate personal networks: Issues and tools for a sociological perspective. Field Methods, 23: 266-86.

Borgatti, S.P., Carley, K.M., and Krackhardt, D. 2006. On the robustness of centrality measures under conditions of imperfect data. Social Networks, 28: 124-36.

Brass, D.J. 1984. Being in the right place: a structural analysis of individual influence in an organization. Administrative Science Quarterly, 29, 518-539.

Brass, D. J., Galaskiewicz, J., Greve, H. R., & Tsai, W. 2004. Taking stock of networks and organizations: A multilevel perspective. Academy of Management Journal, 47: 795-817.

Brewer, D.D. 1995. Patterns in the recall of persons in a department of a formal organization. Journal of Quantitative Anthropology, 5: 255-84.

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

Brewer, D.D. and Webster, C.M. 1999. Forgetting of friends and its effects on measuring friendship networks. Social Networks, 21: 361-73.

Burt, R.S. 1984. Network items and the General Social Survey. Social Networks, 6: 293-339.

Burt, R.S. 1986. A note on sociometric order in the General Social Survey network data. Social Networks, 8, 149-74.

Burt, R.S. 1992. Structural holes. Cambridge, MA: Harvard University Press.

Burt, R.S. 1997. A note on social capital and network constraint. Social Networks, 19: 355-373.

Burt, R.S. 2005. Brokerage and closure. New York: Oxford University Press.

Burt, R.S. 2010. Neighbor networks. New York: Oxford University Press.

Burt, R.S., Kilduff, M. and Tasselli, S. 2013. Social network analysis, through a focus on network advantage. Annual Review of Psychology, 64: In Press.

Coleman, J., Katz, E. and Menzel, H. 1957. The diffusion of an innovation among physicians. Sociometry, 20: 253-70.

Costenbader, E. and Valente, T.W. 2003. The stability of centrality measures when networks are sampled. Social Networks, 25: 283-307.

Feld, S.L. and Carter, W.C. 2002. Detecting measurement bias in respondent reports of personal networks. Social Networks, 24: 365-83.

Fischer, C. S. 1982. To dwell among friends: Personal networks in town and city. Chicago: University of Chicago Press

Golinelli, D., Ryan, G., Green, Jr., H.D., Kennedy, D.P., Tucker, J.S. and Wenzel, S.L., 2010. Sampling to reduce respondent burden in personal network studies and its effect on estimates of structural measures. Field Methods, 22: 217-30.

Hampton, K.N., Sessions, L.F., Her, E.J., and Raine, L. 2009. Social isolation and new technology: How the internet and mobile phones impact Americans' social networks. Pew Internet Report, November 4, 2009, 1-101: http://www.pewinternet.org/~/media//Files/Reports/2009/PIP_Tech_and_Social_Isolation.pdf.

Katz, E. and Lazarsfeld, P.F. 1955. Personal Influence. New York: Free Press.

Killworth, P.D. and Bernard, H.R. 1978. The reverse small-world experiment. Social Networks, 1: 159-192.

Kogovsek, T., Ferligoj, A., Coenders, G. and Saris, W.E. 2002. Estimating the reliability and validity of personal support measures: Full information ML estimation with planned incomplete data. Social Networks, 24: 1-20.

Kogovsek, T., Mrzel, M. and Hlebec, V. 2010. "Please name the first two people you would ask for help": The effect of limitation of the number of alters on network composition. Metodoloski zvezki, 7: 95-106.

Kossinets, G. 2006. Effects of missing data in social networks. Social Networks, 28: 247-68.

Lazarsfeld, P.F., Berelson, B. & Gaudet, H. 1944. The people's choice: How the voter makes up his mind in a presidential campaign. New York: Columbia University Press.

Laumann, E. 1966. Prestige and association in an urban community. New York: Bobbs-Merrill.

Marin, A. 2004. Are respondents more likely to list alters with certain characteristics? Implications for name generator data. Social Networks, 26: 289-307.

Marin, A. and Hampton, K.N. 2007. Simplifying the personal network name generator: Alternatives to traditional multiple and single name generators. Field Methods, 19:163-93.

Marin, A. and Wellman. B. 2010. Social network analysis: An introduction. In Carrington, P. and Scott, J. (Eds.) Handbook of social network analysis. London: Sage.

Marsden, P.V. and Campbell, K.E. 1984. Measuring tie strength. Social Forces, 63: 482-501.

Marsden, P.V. 1987. Core discussion networks of Americans. American Sociological Review, 52:122-131.

Marsden, P.V. 1990. Network data and measurement. Annual Review of Sociology, 16: 435-63.

Marsden, P.V. 2005. Recent developments in network measurement. Pp. 8-30 in Models and Methods in Social Network Analysis, edited by P.J. Carrington, J. Scott, and S. Wasserman. New York: Cambridge University Press.

McCarty, C., Bernard, H.R., Killworth, P.D., Shelley, G.A., and Johnsen, E.C. 1997. Eliciting representative samples of personal networks. Social Networks, 19: 303-23.

McCarty, C., Kilworth, P.D., and Rennell, J. 2007. Impact of methods for reducing respondent burden on personal network structural measures. Social Networks, 29: 300-15.

McPherson, M., Smith-Lovin, L. and Brashears, M.E. 2006. Social isolation in America: Changes in core discussion networks over two decades. American Sociological Review, 71: 353-75.

Merluzzi, J. 2013. Social Capital in Asia: Investigating Returns to Brokerage in Collectivistic National Cultures. Social Science Research, 42: In Press.

Podolny, J.M. and Baron, J.N. 1997. Resources and relationships: Social networks and mobility in the workplace. American Sociological Review, 62: 673-93.

Portes, A. 1998. Social capital: It's origins and applications in modern sociology. Annual Review of Sociology, 24: 1-24.

Schiffer, E. and Hauck, J. 2010. Net-map: Collecting social network data and facilitating network learning through participatory influence network mapping. Field Methods, 22: 231-49.

Tortoriello, M. and Krackhardt, D. 2010. Activating cross-boundary knowledge: the role of Simmelian ties in the generation of innovations. Academy of Management Journal, 53: 167-181.

Tushman, Michael L. (1977). Special boundary roles in the innovation process. Administrative Science Quarterly, 22: 587-605.

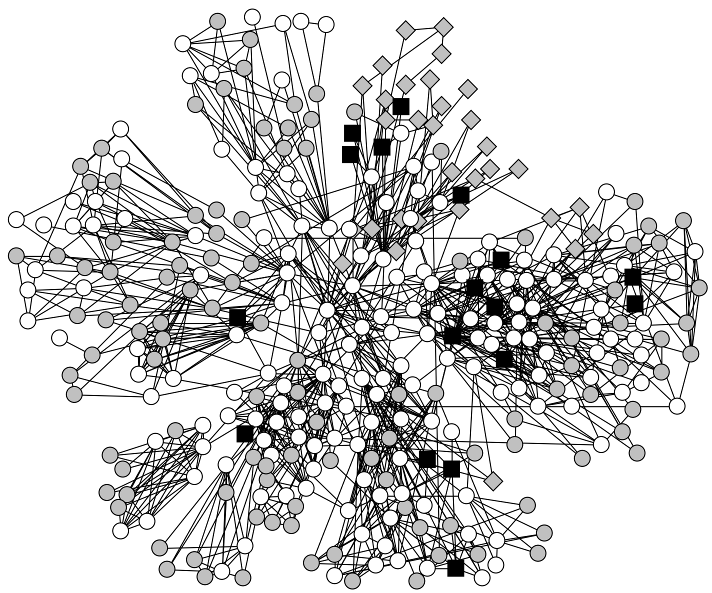
Wellman, B. 1979. The community question: The intimate networks of East Yorkers. American Journal of Sociology, 84: 1201-1231

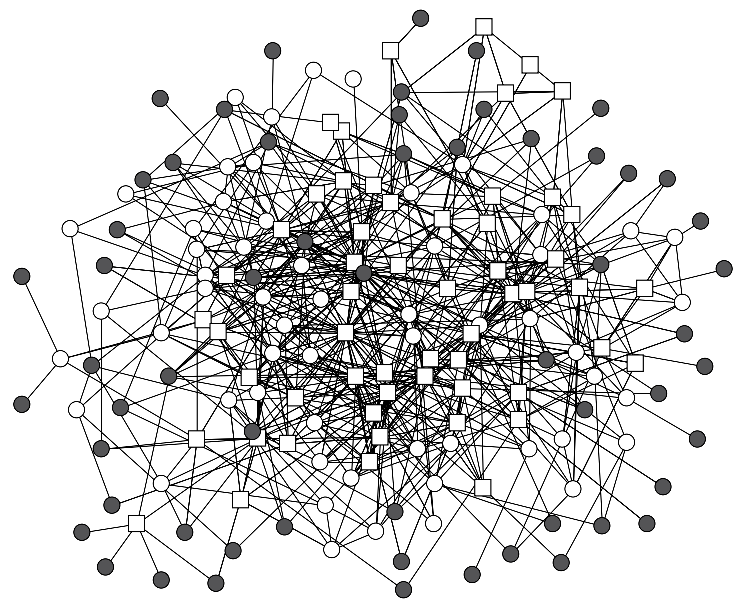
Table 1. Network Metrics on the Two Organizations

		Balkanized Organization	Integrated Organization
Number of Colleagues Cited for Frequent and Substantive Work (Respondents citing each number of contacts)	1	1	4
	2	2	2
	3	4	6
	4	7	6
	5	6	10
	6	13	18
	7	23	30
	8	106	93
Average Number of People Cited by Respondent		7.17 (1,8)	6.95 (1,8)
Average Density among Respondent's Cited Contacts		0.83 (.42,1.00)	0.77 (.51,1.00)
Average Number of Non-Redundant Contacts		2.8 (1,4.50)	2.81 (1,4.40)
Average Years Respondent Has Known Contacts		3.17 (0,30)	5.01 (0,30)
Percent of Cited Contacts Met Daily		46% (0%,100%)	High
Average Number of Days between Meetings with Contacts		8.29 (1,60)	Few
Average Percent of Contacts Outside of Ego's Location		12% (0%, 100%)	11% (0%, 100%)
Average Percent of Contacts Outside Ego's Department		18% (0%, 100%)	68% (0%, 100%)

Note: Parentheses contain the minimum and maximum observed. Averages are based on respondents who cited one or more colleagues for frequent and substantive work.

Figure 1. Frequent and Substantive Work Discussion in Two Organizations.





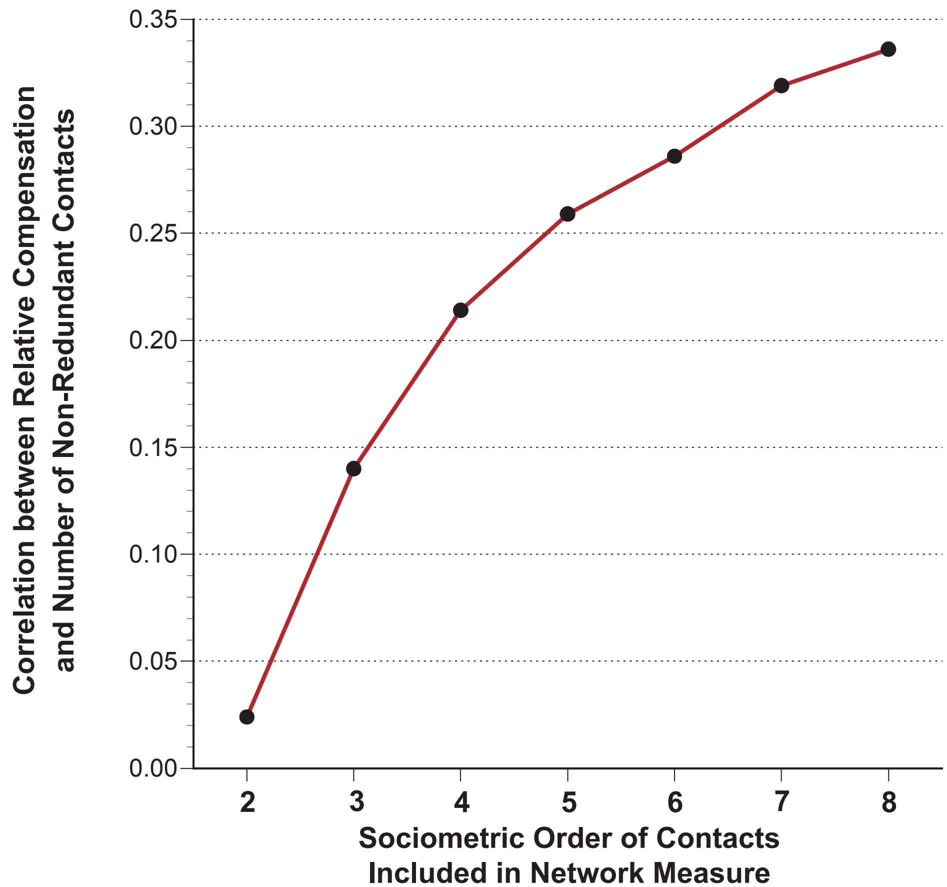
Balkanized Organization

These are 331 managers and senior employees in the Asia-Pacific launch of a company's new software product. Discussion is balkanized within country clusters. White dots indicate people who completed the network survey. Black squares are targeted people who did not respond. Shaded dots are snowball non-responders (diamonds are cited managers outside the study population).

Integrated Organization

These are 113 managers and senior technical people in a software company. Close geographic proximity facilitated the integrated center-periphery structure. White indicates the people who completed one of two network surveys separated by six months. Squares indicate people who completed both surveys. Shaded dots indicate people who did not respond to either survey.

Figure 2. Increasing Association between Compensation and Network Advantage Recording Additional Contacts.



Network Correlation with Compensation (in graph)	.02	.14	.21	.26	.29	.32	.34
Balkanized Organization Only	-.01	.17	.25	.28	.29	.33	.33
Integrated Organization Only	.06	.11	.17	.26	.31	.34	.38
Predicted Compensation (R^2)	.16	.17	.19	.20	.21	.22	.22
Job Rank (t-test)	7.2	6.9	6.7	6.6	6.4	6.2	6.1
Network (t-test)	-0.1	1.6	2.7	3.4	3.7	4.2	4.5

Figure 3. Contact Redundance by Sociometric Order.

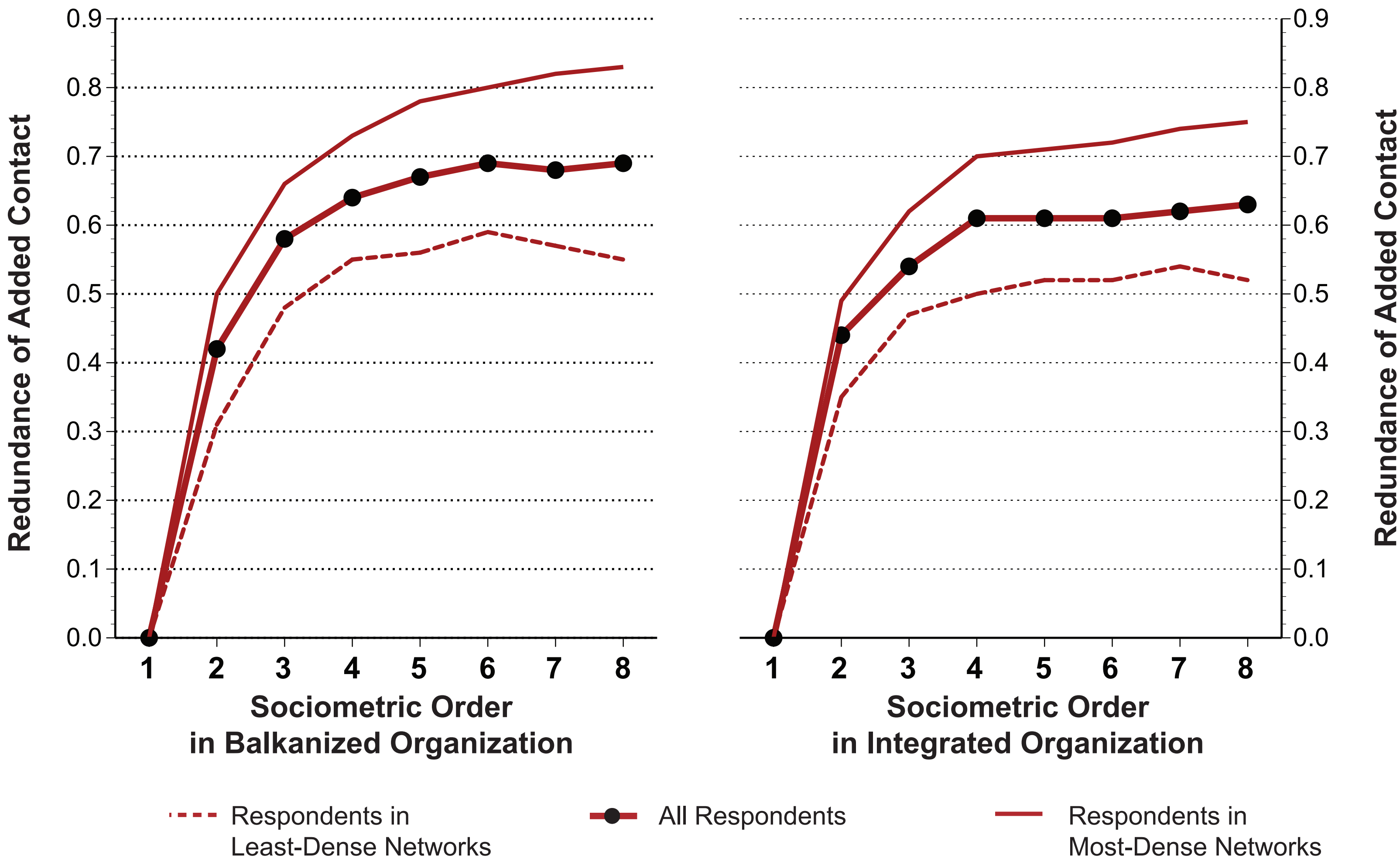


Figure 4. Citations Within Respondent's Department by Sociometric Order.