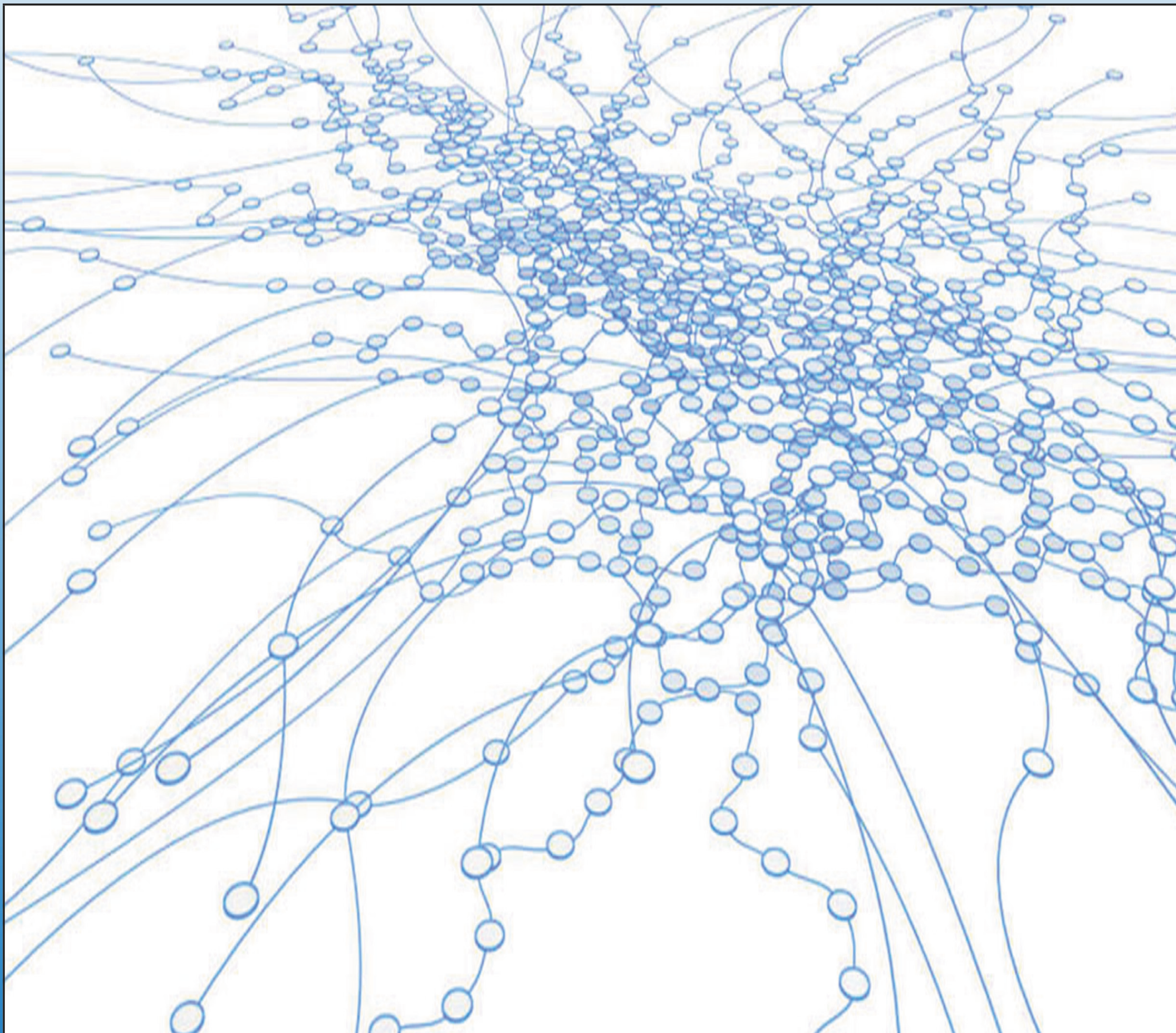
# Connections

Volume 32, Issue 1

June 2012



Official Journal of the International Network for Social Network Analysts

International Network for Social Network Analysis

**Connections** publishes original empirical, theoretical, and methodological articles, as well as critical reviews dealing with applications of social network analysis. The research spans many disciplines and domains including Anthropology, Physics, Sociology, Psychology, Communication, Economics, Mathematics, Organizational Behavior, Knowledge Management, Marketing, Social Psychology, Public Health, Medicine, Computer Science, and Policy. As the official journal for the International Network for Social Network Analysis, the emphasis of the publication is to reflect the ever-growing and continually expanding community of scholars using network analytic techniques. Connections also provides an outlet for sharing news about social network concepts and techniques and new tools for research.

Printing by Corporate Graphics International

**From the Editor**

*Welcome to a new issue of Connections.*

*With the aid of a reconstituted editorial board I have identified a number of target areas for articles. As you can see a number of stylistic changes have been made to make the journal more accessible and to enhance its signature. At the same time, all back issues are being scanned and uploaded so that a complete, open access, record of the journal will be available online. All this will improve the visibility of the association and the value of published articles to our members.*

*My target is to produce two issues a year, and to provide a platform for articles that communicate innovation in the use of methodology and novel SNA applications. I am particularly keen on concise and theoretically embedded articles accessible to an interdisciplinary audience.*

*In this issue we launch a section we call Data Exchange Network (DEN) jointly edited by Rich DeJordy and Pacey Foster. DEN solicits articles explaining data collection, aiming to scrutinize research operationalization. This will promote the wider distribution and use of datasets for teaching and research, but also provide examples of good practice in data collection.*

*To thank our reviewers and authors and to facilitate discussion on the goals of the journal, Rebecca Davis organized for us a very well attended reception at the last Sunbelt. We received ideas and feedback on our plans from a good number of former editors and INSNA members. We look forward to a similar event at the Hamburg Sunbelt in 2013.*

*Concluding, I would like to thank Tom Valente, the preceding editor, and the production team led by Kate Coronges, our Managing Editor, for their support and guidance.*

*I look forward to your contributions in developing the journal.*

*Dimitris*

Dimitris Christopoulos
Editor, Connections
dimitriscc.wordpress.com


## Contents

# How International Are International Congresses?

**Christian Stegbauer**
*Goethe – Universität, Frankfurt am Main, Germany*

**Alexander Rausch**
*Goethe – Universität, Frankfurt am Main, Germany*

**Abstract**
Our study pursues two goals: to present a new method for the analysis of weighted bimodal networks and to show that world congresses lead to fewer international contacts among the contributors than is generally assumed. The study shows that this tendency to endogamy can be observed in the contributors of international congresses. For this purpose two world congresses in the field of sociology are analyzed: the world congress of the IIS in Stockholm (2005) and that of the ISA in Durban (2006). Proceeding from data about the home countries of the contributors in the diverse sessions, a weighted, bimodal network is developed by entering the number of contributors from all the different countries of origin for each session. An analysis of this network represents the focal point of this study. In this context the maximum number of (possible) relationships of attendees from one and the same country is of special interest. These quantities are subjected to a statistical analysis by comparing them to analogously calculated quantities obtained from 1000 randomly drawn bimodal networks with the same marginals as those under discussion. It is found that the homogeneity of the geographical origin of the contributors within the sessions of an international congress is much greater than would be expected by pure coincidence. This holds true even without taking into account the fact that co-authors often come from the same country.

**Authors**
*Christian Stegbauer* is currently a lecturer in sociology at the University of Frankfurt. Research interests include investigating the fundamental of sociology with the aid of network analysis; communications sociology and cultural sociology.

*Alexander Rausch* studied mathematics and physics and is currently working as a senior consultant at the Computing Center of the University of Frankfurt. Teaching: Lectures in Social Network Analysis. Research interests include formal sociology, positional analysis, and analysis of bimodal networks.

*Please address correspondence to Christian Stegbauer at stegbauer@soz.uni-frankfurt.de*

## 1. Introduction

For some time, a trend has been perceived in the sciences, which can be described as denationalization (Crawford et al., 1993). It is said that this process is transnational. The EU is trying to establish a common ground for European science. For this reason, billions of euros are being invested in science programs. National research foundations, e.g. the DFG (German Research Foundation), are supporting visits to international conferences. To become a professor, the candidate should have experience abroad, or at least a publications record of articles in international journals. In Germany, this means in English-language journals. Such considerations are not only fed by a belief in the broader range of English publications, they are justified by the principles of the universality of science. Principles of this universality are "freedom of movement, association, expression and communication for scientists as well as equitable access to data, information and research materials" (International Council for Science, 2004). They also include the rejection of discrimination on the basis of ethnic origin, religion, citizenship, language, political affiliation, gender, sex or age. One reason stated is that, confined to national boundaries, excellent research will not be able to survive – the only way to be successful is through international cooperation (Gibbons et al., 1994) among the most reputed scientists. Furthermore, the need for cooperation is the outcome of the process of functional differentiation (Luhmann, 1977). This process is most visible in the production of industrial goods which are made under conditions of a worldwide division of labor. Like production, all branches of science encompass many subdivisions, each with its own development of methods and theories.

Indeed, there is an international scientific system with publishing bodies, and the number of international co-authorships has increased considerably from the 1980s onwards (Crawford et al., 1993, pg. 4). However, at least in the major countries involved, there co-exists a national system with its own journals and scientific organizations alongside. Careers, too, are often more national than international. Such boundaries can be considered as political borders which are dependent on different political strategies, economic interests and elected governments with budget sovereignty. Such political boundaries and the universality of science are contradictions and international scientific congresses are organized to overcome the boundaries (Hunsaker, 1947).

Observations from the history of science have shown, for instance in sociology, the emergence of a number of theoretical traditions dependent on the cultural background of the societies where they arise (Ekeh, 1974). In addition to propositions which concern the science itself, different styles of scientific debate have emerged (Galtung, 1981). The universality of science is often considered critically, such as being Euro-centric or as perspectives on the West (or should one say the "North"?) (Harding, 1994).

Thus, two tendencies can be seen: the first is a strong trend towards internationalization of the science system; the second, by contrast, the different theoretical approaches and academic styles. This raises the question of how international science is. This issue was investigated taking the example of two world congresses organized by the big sociological associations.

The Institution World Congress has been explored, for example, by Merritt & Hanson (1989). Richard Merritt and Elisabeth Hanson presented a functional analysis of the 11th World Congress of the International Political Science Association held in 1979 in Moscow. They also raised the question of internationality. They used questionnaires for data collection and examined the data by means of regression analysis.International scientific congresses are said to facilitate the formation of transnational networks among scientists. This is one of the main arguments for hosting such large meetings.

On the other hand, experience shows that groups of participants evolve which are relatively homogeneous in terms of country of origin, native language or lingua franca. This also holds, although to a lesser extent, for the contributors to the different sessions of such congresses. At international congresses, it can often be seen that discussions and get-togethers mostly occur within groupings of people who come from the same country or who at least share a common mother tongue.

Merritt and Hanson concluded "that participants were most likely to meet informally with colleagues from their own country, and least likely to enjoy informal contacts with Soviet and East European scholars" (Merritt & Hanson, 1989, pg. 88) The data from the year 1989 were collated in the context of the Cold War, therefore it can be assumed that some changes may have occurred over the last 30 years.

## 2. International Congresses as Bimodal Networks

Contrary to Merritt & Hanson's approach, our investigations drew only on the data published by the organizers of the respective congresses. A further difference is our approach to the analysis of the data, which is based on considerations of network theory.

"Endogamy" is the term used in the present study for the tendency of participants to meet people from the same country in the same session of an international congress. The term "endogamy" seems to be very close to the more usual term of "homophily" (McPherson et al., 2001). McPherson and colleagues define homophily as "the principle that a contact between similar people occurs at a higher rate than among dissimilar people." (McPherson et al. 2001, pg. 416) In the light of this concept, every similarity of adjacent network actors can be interpreted as homophily. We do not want to overstretch the term "homophily". It seems to us that the terms "endogamy" and "exogamy" are more adequate to our subject than the terms "homophily" and "heterophily" because the concept of endogamy and exogamy characterizes the result of a specific selection process. It is widely used in contexts that are similar to that considered in our study. The most comparable use of the term endogamy is by Burris (2004) for the analysis of the academic caste system. "Caste system" stands for different mechanisms of social closure. Endogamy is also used in the field of economics (Carruthers, 1994; Trapido and Hillman, 2010) and more widely in the field of bibliometrics. The term is used to describe the tendency of people to co-author with others from the same country, the same department, the same discipline and so on (López et al., 2010; Lemercier, 2011; Cronin and Meho, 2008; Harnad, 2007; Shadbold, 2006; Bourret et al., 2006; Vázquez-Cupero, 2001).

In this paper an attempt is made to define a measure by which this tendency towards endogamy can be quantified. Furthermore the strength of this effect, in contrast to chance mingling of the participants of a congress, is analyzed. This research approach can be generally applied to bimodal networks and therefore it is relevant beyond the empirical question discussed here.

When measuring the internationality of an international congress, it cannot be determined which participants visited which session, i.e. which participants had a chance to establish informal contact with whom during a session.It is much easier to ascertain who made a previously announced, registered contribution to a session. This is published in the conference program and it includes the country of origin. In the following, only the contributors previously announced in the conference program are considered as the actors of the social network in question. The aim of our study is to analyze the international mixing of these actors aggregated over all sessions of the congress.

To investigate the internationality of the compilation of the different sessions, the following data are required:

- session (title, research committee, etc.)
- contributors (co-authors included)
- contributors' country of origin.

This information is given in the conference program (see Figure 1). A significant endogamy effect among the contributors is evaluated as an indication of a general "tendency towards endogamy" at international congresses.

The analysis is not concerned with the individual actor, but solely with his or her country of origin. For each session, the number of actors from the different countries of origin is noted and thus a bimodal network M is acquired. Figure 2 shows how the data in the congress program are transformed into relational data. Only one country of origin is considered per participant. If a participant has two countries listed, only the first is considered. In our example the result is one participant from the UK, Germany, Russia and Portugal and two participants from France and Spain.

The congress network $\mathcal{M}$ is bimodal because it links two types of objects: the sessions and the country of origin. The network is defined as:

$$\mathcal{M} = \left(m_{i,j}\right)$$

with

$$m_{i,j}$$

**Session** $i = 1,...,s$

**Country of Origin** $j = 1,...,t$

As can be seen, mode 1, the row mode, represents the session and mode 2, the column mode, denotes the country of origin. The network $\mathcal{M}$ is weighted because not only the contributors' countries of origin are recorded, but also the number of contributors from the respective countries. For each row of the weighted bimodal network, the maximum number of symmetric ties among actors from the respective countries can be determined. In this paper, these ties are referred to as *handshakes* and the network generated by these ties a *handshake network*. This term was derived from the fact that contributors of a session usually introduce themselves or are introduced to each other, thus at least affording them an opportunity to shake hands on that occasion.

Sociology and Demography

Chair: John MacInnes, University of Edinburgh (UK) & Autonomous University of Barcelona (Spain)

Room: 354

*Birth Decline and Welfare States. A Comparative Perspective*
Wolfgang Walter, University of Würzburg (Germany)

*Registered Same-Sex Partnership: A Multi-Disciplinary Approach*
Marie Digoix, National Institute for Demographic Studies (France) & Patrick Festy, University of Paris (France)

*The Reproductive Revolution*
Julio Peréz Diaz, Autonomous University of Barcelona (Spain) & John MacInnes, University of Edinburgh (UK) & Autonomous University of Barcelona (Spain)

*Orientation Of Siberian Youth Towards Marriage And Parenthood*
Anna Mikheeva, Institute of Economics and Engineering, Siberian Branch of RAS (Russia)

*Blended Families in Portugal: An Extensive Analysis*
Susana Atalaia, University of Lisbon (Portugal)

**Distributive Papers:**

*Moroccan Inhabitants in Galicia*
Montse Golias, University of Corunna (Spain)

*Figure 1.* Detail of the conference program

| Actors | Countries of Origin |
|---|---|
| John MacInnes | UK ~~& Spain~~ |
| Wolfgang Walter | Germany |
| Marie Digoix | France |
| Patrick Festy | France |
| Julio Peréz Diaz | Spain |
| Anna Mikheeva | Russia |
| Susana Atalaia | Portugal |
| Montse Golias | Spain |

| UK | Germany | France | Spain | Russia | Portugal |
|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 1 | 1 |

Row of weighted bimodal network (empty cells are omitted here)

*Figure 2.* Congress regarded as a bimodal network

As can be seen in Figure 3, the numbers of the respective intranational handshakes are located on the diagonal. The numbers of the respective transnational handshakes are located in the upper triangle of the matrix. Since the matrix is symmetric, the numbers

**row in matrix of the weighted bimodal network**

| UK | Germany | France | Spain | Russia | Portugal |
|---|---|---|---|---|---|
| 1 | 1 | 2 | 2 | 1 | 1 |

**number of persons:** 8
**max. number of handshakes:** 28

**matrix of corresponding *handshake-network***

| | UK | Germany | France | Spain | Russia | Portugal |
|---|---|---|---|---|---|---|
| UK | 0 | 1 | 2 | 2 | 1 | 1 |
| Germany | 1 | 0 | 2 | 2 | 1 | 1 |
| France | 2 | 2 | 1 | 4 | 2 | 2 |
| Spain | 2 | 2 | 4 | 1 | 2 | 2 |
| Russia | 1 | 1 | 2 | 2 | 0 | 1 |
| Portugal | 1 | 1 | 2 | 2 | 1 | 0 |

**max.number of ...**
**intranational handshakes:** 2
**transnational handshakes:** 26

*Figure 3*. The construction of the "handshake network", $\mathcal{H}$

in the lower triangle of the matrix are redundant.

The handshake networks, generated in the described manner for all the sessions, were aggregated (see Figure 4). The result is the aggregated handshake network $\mathcal{H}$ belonging to the weighted bimodal network $\mathcal{M}$.

$h_{j_1,j_2}$ is defined as the maximum number of handshakes among contributors of country $j_1$ and contributors of country $j_2$, aggregated over all sessions.

A formal definition of the elements of the aggregated handshake network $\mathcal{H}$ is given below. The properties of the aggregated handshake network are quite similar to those of the handshake network of a single session:

$$h_{j_1,j_2} = \left\{ \begin{array}{l} \sum_{i=1}^{s} m_{i,j_1} \cdot m_{i,j_2} \\ \sum_{i=1}^{s} \frac{m_{i,j_1} \cdot (m_{i,j_2} - 1)}{2} \end{array} \right\} \quad for \quad \begin{array}{l} j_1 \neq j_2 \\ j_1 = j_2 \end{array}$$

Except for the diagonal elements, matrix $\mathcal{H}$ is identical to the product of $\mathcal{M}$ transposed and $\mathcal{M}$ ( $\mathcal{M}^T\mathcal{M}$ ). Above all, however, the diagonal elements of $\mathcal{H}$ play an essential role in our analysis. The maximum **number of endogamous ties** (that is of all intranational handshakes among contributors) is equal to the sum of the diagonal elements of matrix $\mathcal{H}$.

The maximum **number of exogamous ties** (that is of all transnational handshakes among contributors) is equal to the sum of the elements in the upper triangular matrix of matrix $\mathcal{H}$. By dividing by the respective number of terms of the sums, the **density of endogamous ties** and the **density of exogamous ties** of the weighted bimodal network $\mathcal{M}$ can be defined. By means of these definitions a methodical framework was acquired for the description of endogamy and exogamy within a weighted bimodal network (see Figure 4).

The data of the two world sociology congresses were analyzed: the 37th IIS World Congress held in Stockholm in 2005 and the 16th ISA World Congress held in Durban in 2006 (see Figure 5). First of all, an astonishingly equal distribution of the proportion of contributors from the different countries of origin was found. Among the contributors, the US and the UK have an obvious preponderance. This alone permits anticipation of a high endogamy value for the networks in question. Hence, it is necessary to consider the effect of the different sizes of the delegations from different counties in the discussion of the results. For Stockholm, 67 countries of origin and 300 sessions were found; for Durban 104 countries of origin and 701 sessions.

The values obtained for the density of endogamous ties and of exogamous ties for the two congress networks under consideration are given in Table 1. The density of endogamous ties for the Stockholm congress is almost 19, for exogamous ties it is only around 2. The same situation is found for the Durban congress which yielded an endogamous density of about 29 and an exogamous density of 2.The result is, therefore, that the density of endogamous ties is 10 to 14 times greater than that of exogamous ties. But does this already constitute a significant deviation from a coincidental mixing of the contributors?

| | country 1 | country 2 | country 3 | .. | .. | country t |
|---|---|---|---|---|---|---|
| country 1 | $h_{1,1}$ | $h_{1,2}$ | $h_{1,3}$ | .. | .. | $h_{1,t}$ |
| country 2 | $h_{2,1}$ | $h_{2,2}$ | $h_{2,3}$ | .. | .. | $h_{2,t}$ |
| country 3 | $h_{3,1}$ | $h_{3,2}$ | $h_{3,3}$ | .. | .. | $h_{3,t}$ |
| : | .. | .. | .. | .. | .. | .. |
| : | .. | .. | .. | .. | .. | .. |
| country t | $h_{t,1}$ | $h_{t,2}$ | $h_{t,3}$ | .. | .. | $h_{t,t}$ |

**Endogamy and Exogamy:**

$$En = \sum_{j=1}^{t} h_{j,j} \qquad Ex = \sum_{j_1=1}^{t-1} \sum_{j_2=j_1+1}^{t} h_{j_1,j_2}$$

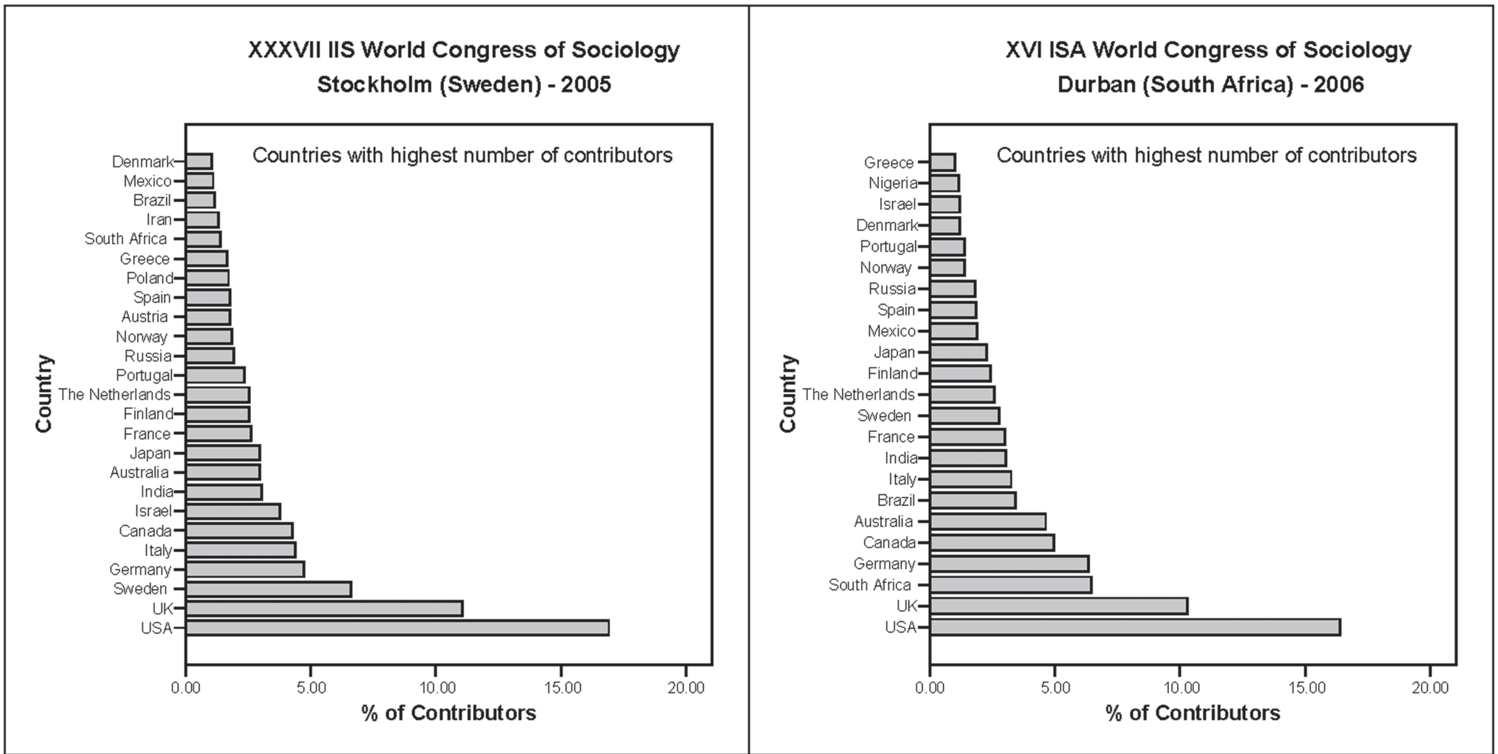
**Associated Densities:**

$$D_{En} = \frac{En}{t} \qquad D_{Ex} = \frac{Ex}{\frac{t(t-1)}{2}}$$

**aggregated maximum number of ...**

- **intranational handshakes En : sum of diagonal**
- **transnational handshakes Ex : sum of upper triangular matrix**

*Figure 4*. Endogamy and exogamy of a weighted bimodal network




| | | | |
|---|---|---|---|
| Number of Countries: | 67 | Number of Countries: | 104 |
| Number of Sessions: | 300 | Number of Sessions: | 701 |

*Figure 5*. The database: IIS & IAS world congresses

**3. A Bimodal Random Network as Test Statistic**

Various approaches to measuring endogamy or homophily can be found in the literature. In contrast to our measure, Burris (2004) gives a heuristic measure based on contingency tables and the respective indifference table. Currarini et al. (2009), by contrast, define a measure reminiscent of Krackhardt's and Stern's E-I-index. But both Burris (2004) and Currarini et al. (2009) focus on the characterization of single groups in the network, not on the network as a whole.

In the present study endogamy is regarded as a property of the network as a whole. In this respect, the approach is quite similar to that of Krackhardt and Stern (1988). Krackhardt and Stern introduced the E-I index to describe the intra- and interdepartmental ties between the employees of a firm. The E-I index is defined as the quotient ( EL - IL ) / ( EL + IL ), where EL is the number of the external (i.e. interdepartmental) links and IL is the number of the internal (i.e. intradepartmental) links in the network analyzed (Krankhardt & Stern 1988:127). Although endogamous links could be regarded as internal links and exogamous links as external links in the sense of Krackhardt & Stern, the E-I index is not adequate for the present research question. A vast surplus of exogamous links over endogamous links is already anticipated and it does not need to be quantified exactly. Instead, the focus of interest is the question to what extent the observed mixing of the participants deviates from coincidental mixing, i.e. whether or not the observed proportion of endogamous links can be explained by chance alone. To answer this question the density of the endogamous ties is taken as a test statistic and the distribution of this parameter is generated by bootstrapping.

For this, the marginal distributions of the weighted bimodal network are considered:

- number and size of sessions
- number of opportunities available to participants from the different countries of origin to submit at least one contribution to a session as the fixed framework for the respective congress.

**Table 1**. Density of endogamous and exogamous ties

| IIS World Congress - Stockholm - 2005 | ISA World Congress - Durban - 2006 |
|---|---|
| $D_{en}$ = 18.82 $D_{ex}$ = 1.89 | $D_{en}$ = 29.21 $D_{ex}$ = 2.06 |
| overall density $D$ = 2.39 | overall density $D$ = 2.63 |

The rationale was to draw bimodal random networks with the same marginal distributions as the observed bimodal network of the respective congress. For these random networks, the associated handshake network and the density of endogamous ties for each case are computed (see Figure 6). The goal is to compare the observed value with the distribution of values gained from the random sampling.

Fixed marginal distributions

| | country 1 | country 2 | country 3 | .. | country j | .. | country t | row marginal |
|---|---|---|---|---|---|---|---|---|
| session 1 | $m_{1,1}$ | $m_{1,2}$ | $m_{1,3}$ | .. | $m_{1,j}$ | .. | $m_{1,t}$ | $m_{1,\bullet}$ |
| session 2 | $m_{2,1}$ | $m_{2,2}$ | $m_{2,3}$ | .. | $m_{2,j}$ | .. | $m_{2,t}$ | $m_{2,\bullet}$ |
| session 3 | $m_{3,1}$ | $m_{3,2}$ | $m_{3,3}$ | .. | $m_{3,j}$ | .. | $m_{3,t}$ | $m_{3,\bullet}$ |
| : | .. | .. | .. | .. | .. | .. | .. | : |
| session i | $m_{i,1}$ | $m_{i,2}$ | $m_{i,3}$ | .. | $m_{i,j}$ | .. | $m_{i,t}$ | $m_{i,\bullet}$ |
| : | .. | .. | .. | .. | .. | .. | .. | : |
| session s | $m_{s,1}$ | $m_{s,2}$ | $m_{s,3}$ | .. | $m_{s,j}$ | .. | $m_{s,t}$ | $m_{s,\bullet}$ |
| column marginal | $m_{\bullet,1}$ | $m_{\bullet,2}$ | $m_{\bullet,3}$ | .. | $m_{\bullet,j}$ | .. | $m_{\bullet,t}$ | |

**Our idea:**
Bimodal random networks are drawn with the same marginal distributions as the observed bimodal network M.
Then the density of the endogamous ties is computed for each random network. Subsequently, the observed value of the density is compared with the simulated random distribution.

*Figure 6*. The construction of the bimodal random graph

The random network obtained has one advantage over conventional theoretical test statistics like the chi-square distribution: it is derived from the collected data. The test statistic can be specifically adapted to our model. For these reasons 1000 simulations were performed in the described manner. The results are shown in Figure 7. The random sampling yields the following distribution of the density of endogamous ties for the respective congresses: both cases reveal that the observed density of endogamous ties is much higher than 99.9% of the respective values obtained by random sampling. Thus it can be inferred that the observed densities deviate significantly from a coincidental mixing of the contributors.

*3.1. Second Model Without Co-Authors*

The result seems clear. There is a marked tendency for participants to meet scientists of their own nationality among the active participants in the sessions. However, there are two objections which demand further research. Firstly, many contributions are not single-authored. In many cases, the presentations are the result of teamwork. Thus, a large number of papers authored by two persons can be found. In some cases more than three scientists are responsible for the papers presented. In many of these cases only one person of the team is present - often this person is the first author (assumption in the second model).

The second objection is that the countries delegate different numbers of scientists. It might now be claimed that a large portion of the observed densities of endogamous ties are due either to the fact that contributions are frequently presented by more than one author of the same country of origin or to the different sizes of delegations. The chance of meeting someone from their own country is dependent on the number of scientists from the same country (especially those from the US and UK). How can these objections be dealt with? In the next step the influence of co-authorship is eliminated. In the second model this effect is suppressed by counting each country of origin only once per contribution (Model 2).

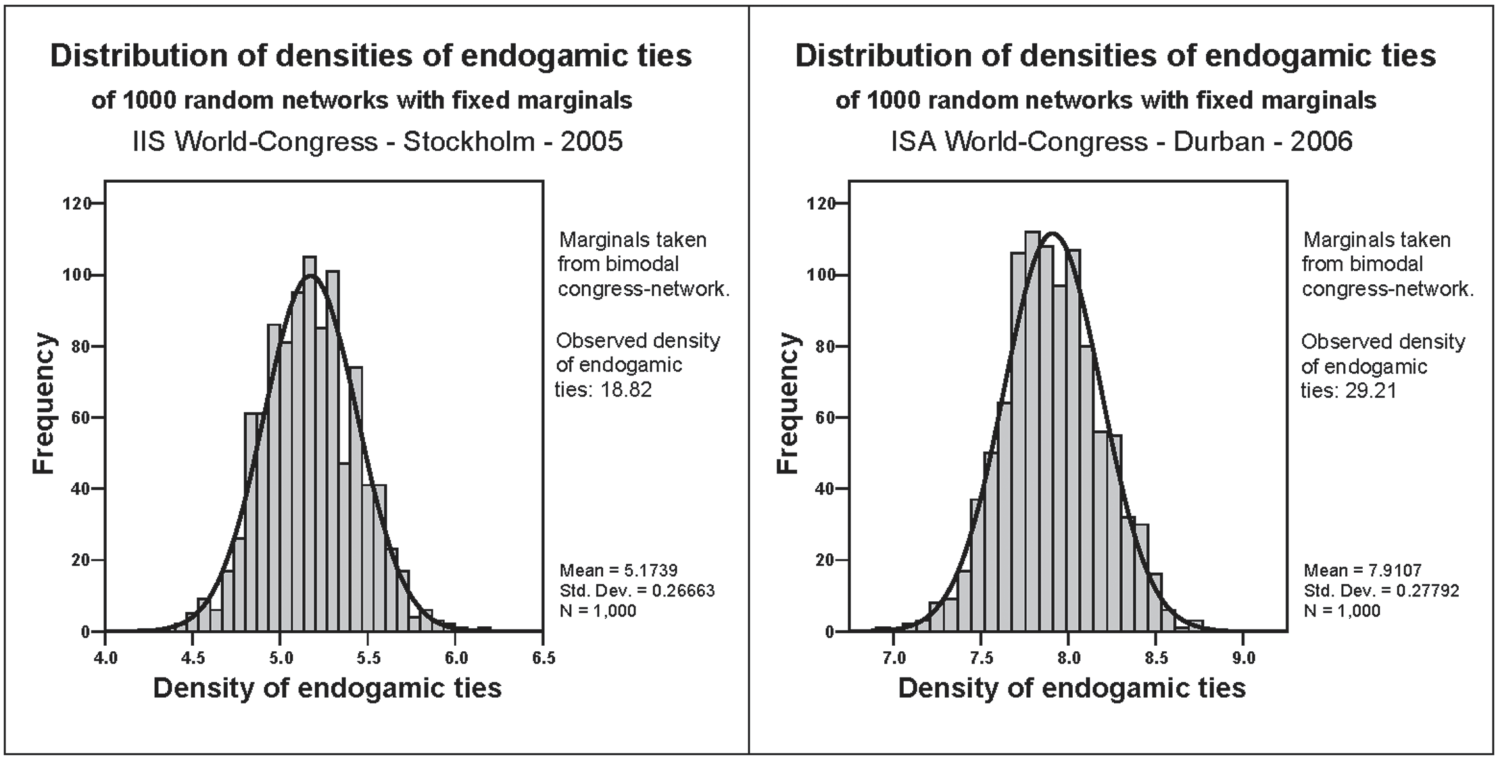
In Figure 8, the original model is compared with Model 2, in which co-authorship is examined. Although the density of endogamous ties is lower in Model 2 (18.82 versus 11.30), the observed value is much higher than 99.9% of the respective values obtained by random sampling. As can be seen, basically this has no impact on the presented result. Hence, it can be maintained that the tendency towards endogamy is not due to co-authorship.

*3.2. Influence of Delegation Size*

Regarding the second objection, in order to eliminate the influence of the size of delegations, the statistical analysis has to be refined. In order to deal with the second objection concerning the size of delegations from different countries, a matrix is presented (see Figures 9 & 10). The elements in the matrix are black if the number of handshakes based on the observed bimodal network is greater than 99.9% of the respective values obtained from the random sample. This shows that significantly increased values are concentrated on the diagonal. Thus, the number of endogamous ties is enhanced in nearly all countries, and this speaks for the assumption that the observed effect is not due to the different sizes of the delegations.

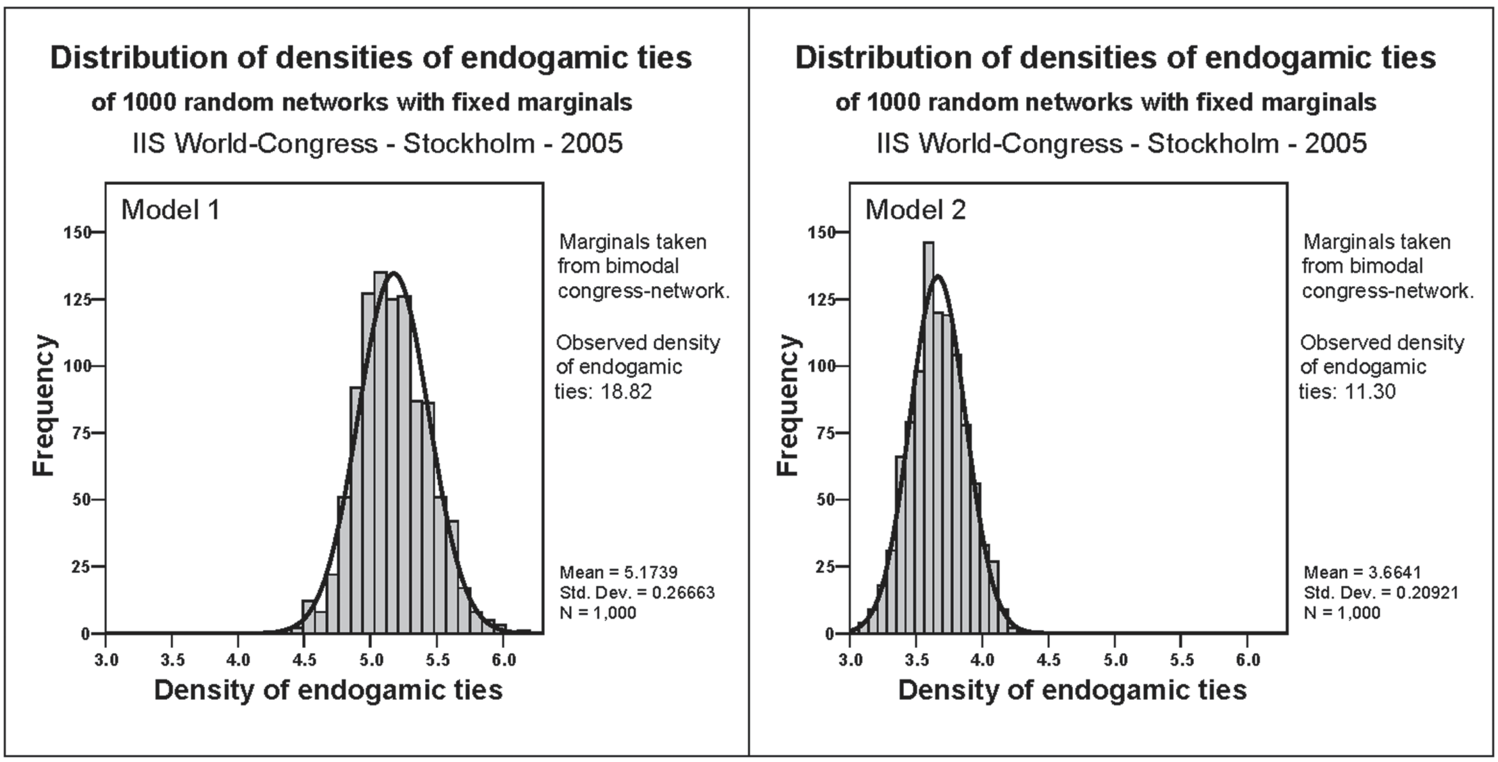
On the other hand, increased values in the upper and lower triangular matrix are distributed more sparsely. Many of these contacts make sense. Hence, increased values are found between neighboring countries, former colonies or between culturally closely connected countries. Examples include the relationships between Greece and Cyprus; Bangladesh, India and Pakistan; Belarus, Hungary, Russia and Slovakia; France and Tunisia; Lithuania and Estonia, etc. Historical relations can be assumed between these countries and maybe exchange programs for scholars exist as well. Some of the strong relationships found cannot be interpreted in the same way. Such relationships are to be found for combinations of countries where at least one of the delegations is very small. This is interpreted as "noise", which is due to the integer characteristic of the problem.

The results for Stockholm can be compared by examining the Durban conference in Figure 10. The situation has not changed. The concentration of endogamous ties is not dependent on the size of delegations. Many of the ties outside the diagonal are expected, but some of them are interpreted as noise.



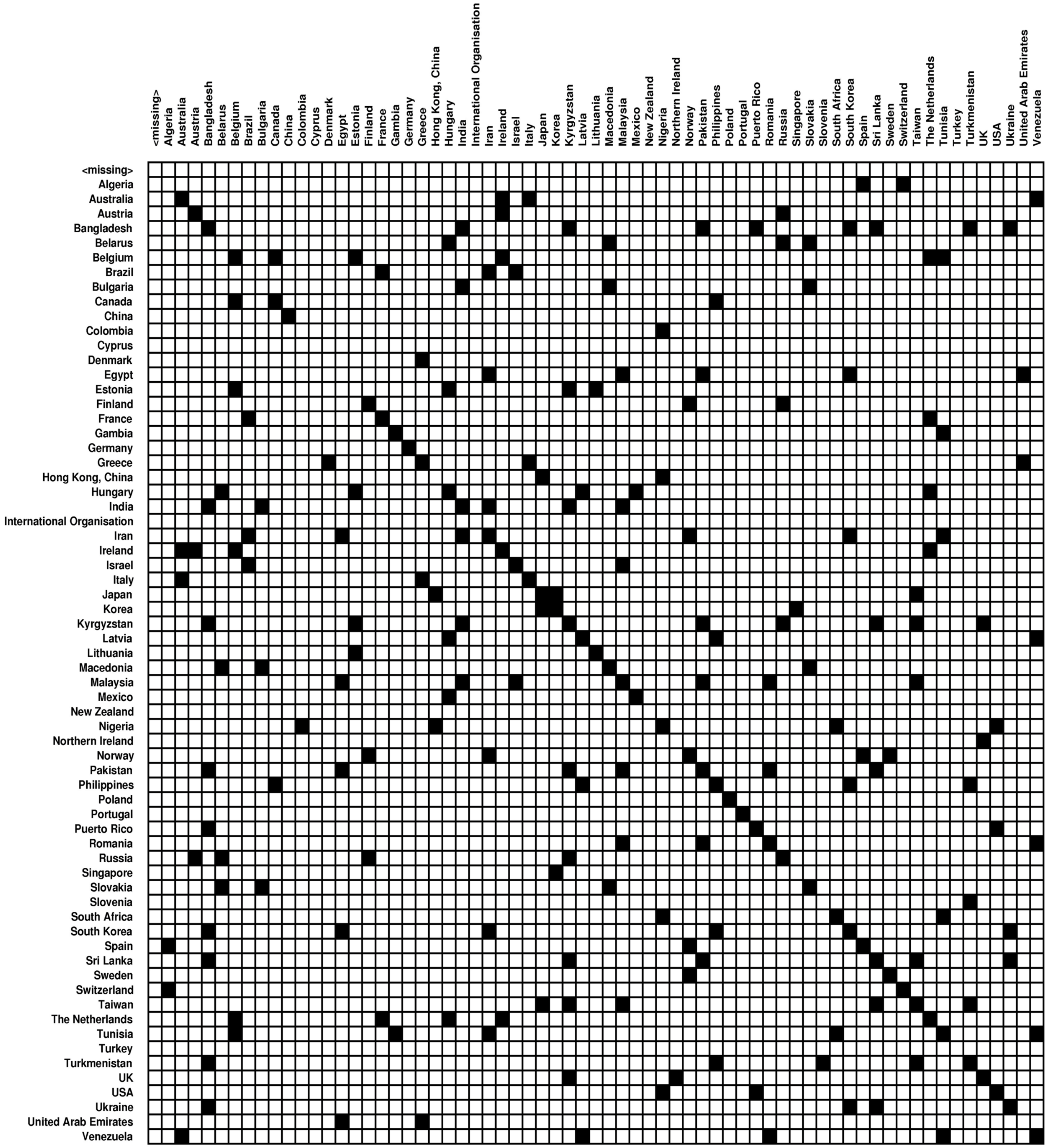

The observed density of the endogamous ties is much higher than 99.9% of the respective values obtained by random sampling.

*Figure 7*. Results of statistical analysis




The effect of co-authorship is suppressed by counting each country of origin only once per contribution (Model 2). This has no impact on the presented result.

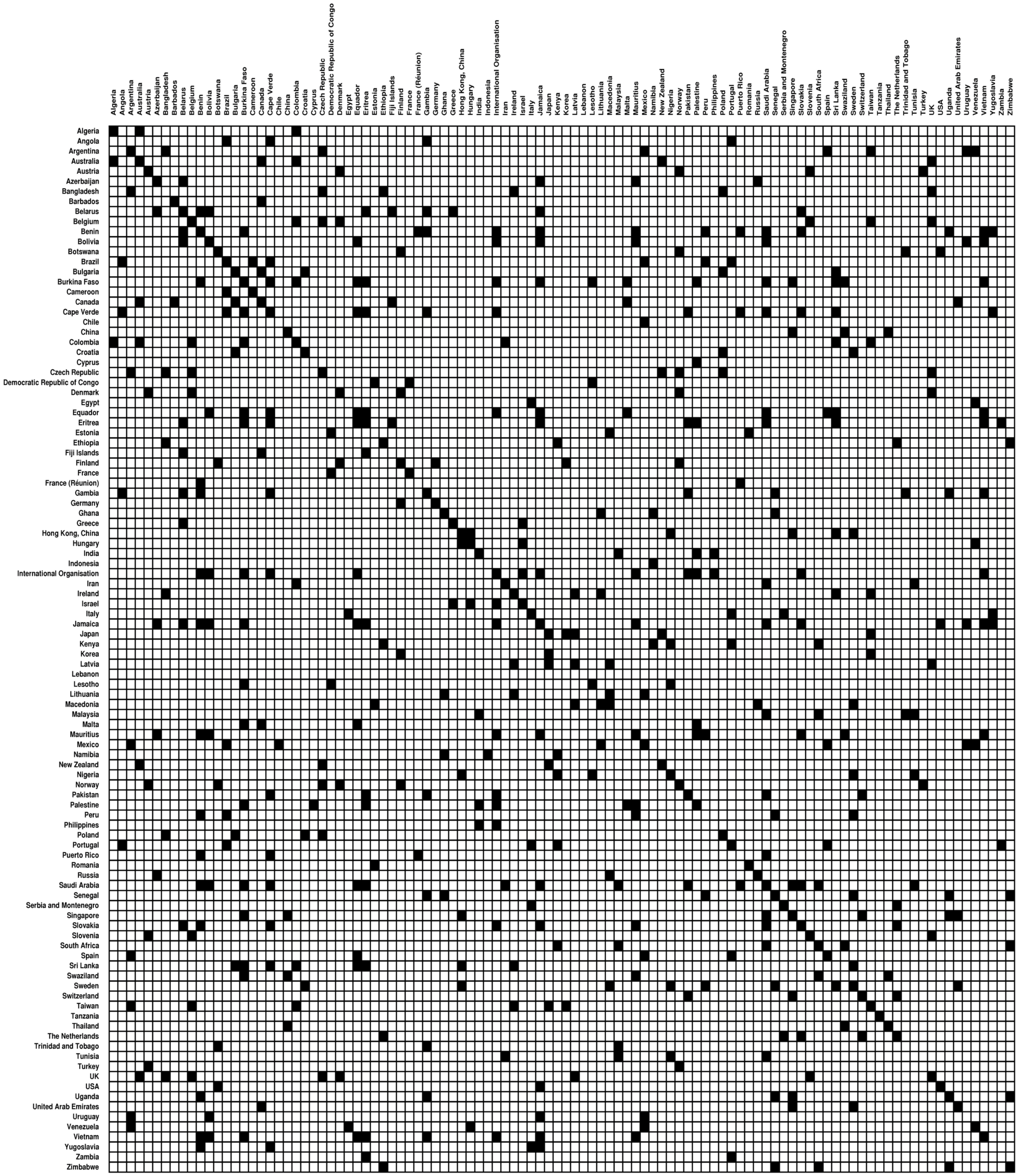
*Figure 8*. Influence of co-authorship




## Notes

- Elements are marked black if the number of handshakes based on the observed bimodal network is greater than 99.9% of the respective values gained from the random sample.
- The number of endogamous ties is higher for nearly all countries.
- This speaks for the assumption that the observed effect is not due to the different sizes of the delegations.

*Figure 9*. Influence of size of delegations (Stockholm)




## Notes

- Elements are marked in black if the number of handshakes based on the observed bimodal network is greater than 99.9% of the respective values gained from the random sample.

*Figure 10.* Influence of sizes of delegations (Durban)

## 4. Interpretation

It seems that national boundaries are not easy to overcome. The present investigation did not reveal that sociology is now denationalized. The conclusion to be drawn from these empirical data is that the utopia of universal sciences is not reflected by reality. The data do, however, provide food for thought about the reasons for the results of this study.

In the first place, it may be an effect of the way in which the congresses are organized. Sessions of the congresses are organized by the board of the International Sociological Association or the International Institute of Sociology. The ISA also has research committees with chairpersons and other officials. In many cases they organize the different sessions or some scholars have an idea for a session and submit it when the conference is announced (often with suggestions for potential contributors). Presentations submitted in response to a call-for-papers are evaluated by persons who have been socialized in a specific scientific culture.

On the one hand, organizers have their own personal networks. Regarding the networks of organizers and contributors, it is reasonable to assume that a proposal is more likely to be submitted if the organizer of a session is known personally and it is more likely for the proposal to be accepted under these circumstances.On the other hand, participants import parts of their personal networks into a congress. This bolsters self-confidence in an otherwise unknown environment, and it is generally recognized that it also binds a great amount of time and cognitive resources.

Regardless of this, the different academic styles, research traditions and regional references and relationships could cause a tendency towards endogamy. Johan Galtung (1981) differentiated between Saxonic, Teutonic, Gallic and Japonic styles. Differences can be found with respect to the orientation and weight of theoretical and national traditions. A relationship between cultural traditions and scientific thinking is evident (Ekeh 1974). Hotly debated theories in one country will be less important in another.

The relevance of theories is not only the result of differences in style and cultural background. Their importance changes with social problems, differences in institutions, e.g. educational systems, etc. Not all themes have the same relevance for every country. For instance, the problems of unemployment, migration, health and the welfare state are unequally distributed. For this reason, research programs to discover sociological aspects of these problems differ according to the urgency of these issues. Examples of the differences described are sessions like: “School in the Frontiers of Modernity - the Mediterranean Space” or “Shifting Boundaries of Knowledge: Creating Spaces for Social Science, Law and Humanities in South Africa”.

Three further reasons may play a role. Most members of delegations are foreigners in the environment of the congress. The people they know are, to a great extent, from their own country. Cultural peculiarities of the same provenance can be understood more easily than the differences which arise from cross-cultural characteristics. Last but not least, most academic careers are ‘national’, hence career networks focus on intranational rather than on international contacts.

## References


Bourret, P., Mogoutov, A., Julian-Reynier, C., & Cambrosio, A. (2006). A New Clinical Collective for French Cancer Genetics: A Heterogeneous Mapping Analysis. *Science Technology Human Values, 31(4)*, 431-464. doi: 10.1177/0162243906287545

Burris, V. (2004). The Academic Caste System: Prestige Hierarchies in PhD Exchange Networks. *American Sociological Review, 69(2)*, 239-264.

Currarini, S., Jackson, M. O., & Pini, P. (2009). An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica, 77(4)*, 1003-1045.

Carruthers, B. G. (1994). Homo Economicus and Homo Politicus: Non-Economic Rationality in the Early 18th Century London Stock Market. *Acta Sociologica, 37(2)*, 165-194.

Crawford, E. T., Shinn, T., & Sörlin, S. (1993). The Nationalization and Denationalization of Science. An Introductory Essay. In E. T. Crawford, T. Shinn, & S. Sörlin (Eds.): Denationalizing science. The contexts of international scientific practice (Sociology of the Sciences, Vol. 16). (pp. 1-42) Dordrecht: Kluwer.

Cronin, B., & Meho, L. I. (2008). The shifting balance of intellectual trade in information studies. *Journal of the American Society for Information Science and Technology, 59(4)*, 551-564.

Ekeh, P. P. (1974). Social exchange theory. The two traditions. London: Heinemann.

Galtung, J. (1981). Structure, Culture and Intellectual Style: An Essay Comparing Saxonic, Teutonic, Gallic and Nipponic Approaches. *Social Science Information, 20(6)*, 817-856.

Gibbons, M., Limonges, C., Schwartzman, S., Scott, P., Trow, M. & Nowotny, H., (1994). The New Production of Knowledge. The Dynamics of Science and Research in Contemporary Societies. London: SAGE.

Harnad, S. (2007). Open Access Scientometrics and the UK Research Assessment Exercise. Proceedings of 11th Annual Meeting of the *International Society for Scientometrics and Informetrics 11(1)*, 27-33. Retrieved from http://eprints.ecs.soton.ac.uk/13804/

Harding, S. G. (1994). Is Science Multicultural?: Challenges, Resources, Opportunities, Uncertainties. *Configurations, 2(2)*, 301-330.

Hunsaker, J. C. (1947). International Scientific Congresses. Proceedings of the *American Philosophical Society, 91(1)*, 126-128.

International Council for Science. (2004). Universality of Science in a Changing World. Retrieved from http://www.icsu.org/Gestion/img/ICSU_DOC_DOWNLOAD/567_DD_FILE_Universality.pdf

Krackhardt, D. (1987). Cognitive Social Structures. *Social Networks, 9(2)*, 109-134.

Krackhardt, D. & Stern, R. N. (1988). Informal Networks and Organizational Crises: An Experimental Simulation. *Social Psychology Quarterly, 51(2)*, 123-140.

Lemercier, C. (2011). L'Académie des sciences vue de l'intérieur: plus de discussions que de tirs, décidément. Retrieved from http://evaluation.hypotheses.org/1164

López, L. W.; Garcia-Cepero, M. C, Bustamente, A., Constanza, M., Silva, L. M., & Aguado López, E. (2010). General Overview of Academic Production in Ibero-American Psychology, 2005-2007: *Papeles del Psicólogo 31(3)*, 296-309.

Retrieved from http://javeriana.academia.edu/Wilsonlopezlopez/Papers/438377/General_overview_of_academic_production_in_colombian_psychology_indexed_by_psicoredalyc_2005-2007

Luhmann, N. (1977). Differentiation of Society. *Canadian Journal of Sociology, 2(1)*, 29-53.

Merritt, R. L. & Hanson, E. C. (1989). Science, Politics and International Conferences. A Functional Analysis of the Moscow Political Science Congress. Boulder, Colorado: Lynne Rienner.

McPherson, M., Smith-Lovin, L., & Cook J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology, 27(1)*, 415-444.

Shadbolt, N., Brody, T., Carr, L., & Harnad, S. (2006). The Open Research Web: A Preview of the Optimal and the Inevitable. [Book Chapter] (In Press) Retrieved from http://cogprints.org/4841/2/shad%2Dbch.pdf

Trapido, D., & Hillmann, H. (2010). Relational Counterbalances to Economic Endogamy: A Theory and a Historical Example. Centre of Organizational Research (COR), University of California, Irvine. Retrieved from http://www.cor.web.uci.edu/sites/cor.web.uci.edu/files/u3/TrapidoWorkshopPaper.pdf

Vázquez-Cupeiro, S. (2001). A Qualitative Review of the University in Spain – Meritocracy, Endogamy and the Gendered Opportunity Contexts. [Training Paper TP 2003-01]. Retrieved from http://csn.uni-muenster.de/women-eu/download/TP%2003-01%20Vasquez-Cupeiro%20Fernandez.pdf

# Network Topography, Key Players and Terrorist Networks

**Sean F Everton**
*Department of Defense Analysis, Naval Postgraduate School*
*Monterey, California*

**Abstract**
In recent years social network analysis (SNA) has enhanced our understanding of how terrorist networks organize themselves and has offered potential strategies for their disruption. To date, however, SNA research of terrorist networks has tended to focus on key actors within the network who score high in terms of centrality or whose structural location (i.e., their location within the overall network) allows them to broker information and/or resources within the network. However, while such a focus is intuitively appealing and can provide short-term satisfaction, it may be putting the cart before the horse. Before jumping to the identification of key actors, we need to first explore a network's overall topography. Research suggests that networks that are too provincial (i.e., dense, high levels of clustering, an overabundance of strong ties) too cosmopolitan (i.e., sparse, low levels of clustering, an overabundance of weak ties), too hierarchical (i.e., centralized, low levels of variance) and/or too heterarchical (i.e., decentralized, high levels of variance) tend not to perform as well as networks that maintain a balance between these extremes. If these dynamics hold true for terrorist networks as well, then the key player approach may be appropriate in some circumstances, but may lead to deleterious results in others. More importantly, it suggests that analysts need to consider a network's overall topography before crafting strategies for their disruption.

**Authors**
*Sean F Everton, PhD* is an Assistant Professor in the Defense Analysis Department at the Naval Postgraduate School in Monterey, CA. He is also the Co-Director of the Common Operational Research Environment (CORE) Lab, which is part of the Defense Analysis Department and specializes in the application of new analytical methodologies (e.g., social network and geospatial analysis) to the crafting of strategies at the operational and tactical levels. While most of his work focuses on social network analysis, he has also published in the areas of sociology of religion, economic sociology, and political and social movements.

*Please address correspondence to Sean Everton, PhD, Assistant Professor. Co-Director, CORE Lab Defense Analysis Department, Naval Postgraduate School, Monterey, CA 93943; Email sfeverto@nps.edu*

## 1. Introduction and Background

Roberts and Everton (2009) recently argued that while social network analysis (SNA) has wide appeal as a methodology for targeting members of terrorist networks, it possesses a much wider application than is currently being used. Furthermore, they argued that strategy should drive the choice of metrics rather than the other way around. Unfortunately, just the opposite appears to be happening. The tail (i.e., the choice of metrics) is wagging the dog (strategic choices). Indeed, the most common application of SNA to the study of terrorist networks has been the key player approach, which focuses on targeting key actors within the network for elimination or capture (a.k.a. the "whack-a-mole" strategy).

While the focus on key individuals is intuitively appealing and might provide short-term results, such a focus may be misplaced and, in fact, may make tracking, disrupting and destabilizing terrorist networks more difficult. As Brafman and Beckstrom (2006) have noted, targeting key players in decentralized organizations seldom shuts them down. Instead, it only drives them to become more decentralized, making them even harder to target.

In terms of terrorist networks, such a strategy may in fact exacerbate what Sageman (2008) refers to as the "leaderless jihad," by which he means the numerous independent and local groups that have branded themselves with the Al Qaeda name and are attempting to emulate bin Laden and his followers in conceiving and executing terrorist operations from the bottom up.Here it is suggested that analysts need to first explore a terrorist network's overall topography (i.e., its level of density, centralization, clustering, etc.) before estimating brokerage, centrality and other types of metrics.

This is not to say analysts have completely neglected the topographical dimensions of terrorist networks. There have been exceptions. Pedahzur and Perliger (2006), for example, noted that terrorist networks with a large number of cliques appear to be more effective than those with few, and the U.S. Army's most recent counterinsurgency manual (U.S. Army, 2007) argues that network density is positively associated with network efficiency and, as such, should guide tactics. Perhaps the best known example is Sageman's (2004b) initial study of what he calls the Global Salafi Jihad (GSJ) in which he found that the GSJ exhibits the characteristics of a scale-free network. This discovery led him to argue that the United States should focus its efforts on taking out hubs (i.e., nodes that have many connections) rather than randomly stopping terrorists at borders. "[The latter] may stop terrorists from coming here, but will leave the network undisturbed. However… if the hubs are destroyed, the system breaks down into isolated nodes or sub-groups. The jihad will be incapable of mounting sophisticated large scale operations like the 9/11 attacks and be reduced to small attacks by singletons" (Sageman, 2004a).

While the simultaneous removal of 10-15% of a terrorist network's hubs is easier said than done, and subsequent research has found that hubs are often quickly replaced by other, highly central and/or structurally equivalent actors (Pedahzur & Perliger, 2006; Tsvetovat & Carley, 2005), it does not change the fact that Sageman's approach illustrates how the exploration of a network's overall topography can inform strategic decision-making.

## 2. Hypothesis and Aims

In this paper I explore two interrelated but analytically distinct topographical dimensions of networks that appear to affect network performance: what I call the (1) provincial-cosmopolitan and (2) heterarchical-hierarchical dimensions. I begin by drawing on "light network" research, defined as networks that are overt and legal as opposed to "dark networks," which are covert and illegal networks such as terrorist networks (Milward & Raab, 2006; Raab & Milward, 2003), This research suggests that networks that are too provincial (e.g., dense, high levels of clustering, an overabundance of strong ties) or too cosmopolitan (e.g., sparse, low levels of clustering, an overabundance of weak ties) tend to perform more poorly than networks that maintain a balance between the two.

Next, I turn to a series of studies that suggest that a similar dynamic is at work in terms of how hierarchical a network is: networks that are too hierarchical (e.g., centralized, high levels of variance) or too heterarchical (e.g., decentralized, low levels of variance) tend to under-perform those that lie between the two extremes. I then note that if these same dynamics hold true for terrorist and other forms of dark networks, then the central actor approach may be appropriate in some circumstances but not in others. More broadly I argue that analysts need to take into account a network's overall topography before crafting strategies or their disruption. Consequently, I conclude by suggesting what these studies imply for strategic decision-making in terms of tracking and disrupting dark networks.

## 3. Types of Networks

### *3.1 Provincial and Cosmopolitan Networks*

In what is now regarded as a classic study, Granovetter (1973, 1974) discovered that when it came to finding their current job people were far more likely to have used personal contacts than other means. Moreover, of those who found their jobs through personal contacts, most were weak ties (i.e., acquaintances) rather than strong ones (i.e., close friends). This occurred because people tend to have more weak ties than strong ties (because weak ties demand less of our time), and because weak ties are more likely to form the crucial bridges that tie together densely knit clusters of people (see Figure 1). Granovetter argued that weak ties often connected otherwise disconnected groups. Thus, whatever is to be spread (e.g., information, influence, and other types of resources), it will reach a greater number of people when it passes through weak ties rather than strong ones (Granovetter, 1973, pg. 1366). Moreover, actors with few weak ties are more likely to be "confined to the provincial news and views of their close friends" (Granovetter, 1983:202).

Granovetter does not argue that strong ties are of no value. He notes that while weak ties provide individuals with access to information and resources beyond those available in their immediate social circles, strong ties have greater motivation to be sources of support in times of uncertainty (Granovetter, 1983, pg. 209). Others have noted this as well (see e.g., Krackhardt, 1992; Stark, 2007). "There is a mountain of research showing that people with strong ties are happier and even healthier because in such networks

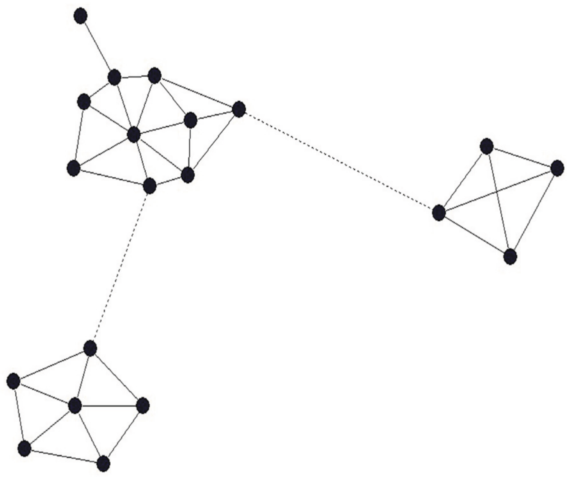


*Figure 1.* Strong and weak ties (Granovetter, 1973, 1983)

members provide one another with strong emotional and material support in times of grief or trouble and someone with whom to share life's joys and triumphs" (Stark, 2007:37). This suggests that people's networks differ in terms of their mix of weak and strong ties. Individuals' networks range from local or provincial ones, consisting of primarily of strong, redundant ties and very few weak ties, to worldly or cosmopolitan ones, consisting of numerous weak ties and very few strong ties (Stark, 2007:37-38). It also suggests that peoples' networks should ideally consist of a mix of weak and strong ties. They should be neither too provincial nor too cosmopolitan but rather land somewhere between the two extremes.

Pescosolido and Georgianna's (1989) study of suicide illustrates this dynamic. It found that social network density has a curvilinear (or inverted U) relationship to suicide. Individuals who are embedded in very sparse (i.e., cosmopolitan) and very dense (i.e., provincial) social networks are far more likely to commit suicide than are people who are embedded in moderately dense networks. Why? People who are embedded in sparse social networks often lack the social and emotional ties that provide them the support they need during times of crisis. They also typically lack ties to others who might otherwise prevent them from engaging in self-destructive (i.e., deviant) behavior (Finke & Stark, 2005; Granovetter, 2005). On the other hand, individuals who are embedded in very dense networks are often cut-off from people outside of their immediate social group, which increases the probability that they will lack the ties to others who could prevent them from taking the final, fatal step.

An ideal mix of weak and strong ties appears to provide benefits at the individual level as well as at the organizational level. In his study of the New York garment industry, Brian Uzzi (1996) found that a mix of weak and strong ties proved beneficial to the long-term survival of apparel firms. The firms he studied tended to divide their market interactions into two types: "market" or "arms-length" relationships (i.e., weak ties) and "special" or "close" relationships (i.e., strong ties), which Uzzi refers to as "embedded" ties. He found that while market ties were more common than embedded ones, the latter tended to be more important in situations where trust was of overriding importance, where detailed information had to be passed to others, and when certain types of joint problem-solving were on the table (Uzzi, 1996:677). According to Uzzi, embeddedness increases economic effectiveness along a number of dimensions crucial to competitiveness in the global economy: organizational learning, risk-sharing and speed-to-market. However, he also found that firms that are too embedded often suffer because they do not have access to information from distant parts of the network, which makes them vulnerable to a rapidly changing environment. This led him to argue that firms should seek to maintain a balance of embedded and market ties and found that an inverted U relationship exists between the degree of embeddedness and the probability of firm failure (Uzzi, 1996:675-676).

Interestingly, Uzzi and Spiro (2005) found that an inverted U relationship also existed in the extent to which the networks of creative teams producing Broadway musicals from 1945 to 1989 exhibited "small-worldness" and the probability that a musical would be a critical and financial success. They believe that this relationship existed because up to a point, connectivity and cohesion facilitate the flow of diverse and innovative material across the network. Moreover, connectivity and cohesion make risk-taking among the teams more likely because they are embedded in networks of trust.

> *As the level of Q increases, separate clusters become more interlinked and linked by persons who know each other. The processes distribute creative material among teams and help to build a cohesive social organization within teams that support risky collaboration around good ideas* (Uzzi & Spiro, 2005:464).

However, as connectivity and cohesion increase, homogenization and imitation set in and returns become negative.

> *Increased structural connectivity reduces some of the creative distinctiveness of clusters, which can homogenize the pool of creative material. At the same time, problems of excessive cohesion can creep in. The ideas most likely to flow can be conventional rather than fresh ideas because of the common information effect and because newcomers find it harder to land "slots" on productions* (Uzzi & Spiro, 2005, pg. 464).

In other words, initially connectivity and cohesion increase a network's overall creativity by encouraging human innovation, but beyond a certain point, they begin to stifle it.

While it may be (morally) difficult to conceive of terrorist networks as varying in their ability to encourage innovative thinking and creative risk-taking, these studies should give us pause. They suggest that in order to be successful, terrorist networks can be neither too provincial nor too cosmopolitan. Of course, what constitutes the optimum balance of strong and weak ties will most likely vary depending on the environment in which it operates (e.g., the IRA can operate more openly in Ireland than Al-Qaeda can in the United States), but that still should not discourage analysts from exploring and documenting this topographical feature of dark networks.

*3.2 Heterarchical and Hierarchical Networks*

Another more well-developed body of research has explored how the degree to which an organization is hierarchically structured impacts its performance (see e.g., Nohria & Eccles, 1992; Podolny & Page, 1998; Powell, 1985, 1990; Powell & Smith-Doerr, 1994). This literature typically identifies two ideal types of organizational form: networks and hierarchies. The former are seen as decentralized, informal and/or organic, while the latter are seen as centralized, formal and/or bureaucratic (Burns & Stalker, 1961; Powell, 1990; Ronfeldt & Arquilla, 2001). While this distinction is useful (and appropriate) in some contexts (see, e.g. Arquilla & Ronfeldt, 2001; Castells, 1996; Podolny & Page, 1998; Powell & Smith-Doerr, 1994; Ronfeldt & Arquilla, 2001), it is probably better to think of these two ideal types as poles on either end of a continuum, running from highly decentralized forms on one end to highly centralized forms on the other.

More importantly, at least for our purposes here, research suggests that this dimension impacts network performance much like the provincial-cosmopolitan dimension: that is, an optimal level of centralization or hierarchy exists. For example, Rodney Stark (1987, 1996), in his analysis of why some new religious movements succeed, identified centralized authority as an important factor. Nevertheless, he notes that too much centralization can be a bad thing and successful religious movements, such as the Mormon (LDS) Church, balance centralized authority structures with decentralized ones:

> *But it would be wrong to stress only the hierarchical nature of LDS authority and its authoritarian aspects, for the Latter-day Saints display an amazing degree of amateur participation at all levels of their formal structure. Moreover, this highly authoritarian body also displays extraordinary levels of participatory democracy—to a considerable extent the rank-and-file Saints are the church. A central aspect of this is that among the Latter-day Saints to be a priest is an unpaid, part-time role that all committed males are expected to fulfill* (Stark, 2005, pg. 125).

Like the provincial-cosmopolitan dimension, the optimal level along the heterarchical-hierarchical dimension varies depending on environmental context. Decentralized structures are generally seen as better suited for solving nonroutine, complex and/or rapidly-changing problems or challenges because of their adaptability, while centralized ones are better suited for stable environments where economies of scale are of paramount importance (Granovetter, 1985; Raab & Milward, 2003). Saxenian (1994, 1996), for instance, contends that Silicon Valley emerged as the center of the high technology universe because it developed a highly-flexible industrial network — characterized by a horizontally integrated industrial system, flat corporate structures, friendly local institutions, a supportive culture and a heterarchical institutional infrastructure — that was more responsive to the volatile high technology industry than were other regional areas. And, in his reflection on the structure of terrorist organizations, David Tucker (2008) argues that while network forms of organization are useful for some tasks (e.g., mobilization), they are not useful for others (e.g., security). He also notes that the optimal form of organization depends largely on the environment in which an organization operates:

> *The most important issue is how well an organization's structure is adapted to its environment, which includes what its enemies are doing, given what the organization wants to achieve and the resources available to it. No one organizational structure is always inherently superior to another. Some are better for some things, some for others. These principles apply to al Qaeda as well as the governmental network (the federal, state, and local governments) in the United States* (Tucker, 2008, pg. 2).

Once again too much of a good thing can lead networks to underperform, and unless it is demonstrated otherwise, there is no reason to suspect that this same dynamic applies to terrorist networks. From their perspective they cannot be too centralized or decentralized, while from ours that is exactly how we want them to be.

**4. Strategic Implications**

This brief analysis of the relationship between network effectiveness and network topography suggests that analysts seeking to disrupt dark networks will want to pursue policies that push dark networks toward the tails of these two dimensions (see Figure 2).

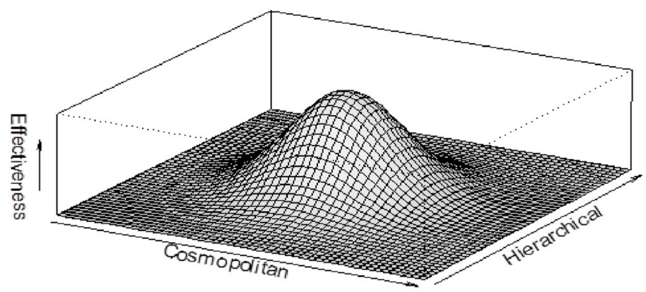



*Figure 2.* Hypothesized relationship between network topography and effectiveness (Note: Graphic generated using R (R Development Core Team, 2009)

For example, a scenario where analysts are seeking to disrupt a terrorist network that lies on the centralized side of the continuum. If, in such a scenario, they target a central actor for capture or elimination and are successful, they may cause the network to become less centralized and actually more effective. Instead, they may want to implement a misinformation campaign that breeds distrust between the network's inner circle and its peripheral members that will hopefully lead the former to centralize decision-making, communication and strategic functions even more than they currently are. Or again, analysts may seek to disrupt a somewhat provincial terrorist network by adopting a strategy that causes it to turn in on itself (e.g., peeling off peripheral members through an amnesty campaign), thus making it more provincial and less effective.

The important point here is that the topographical features of terrorist networks should inform strategic decision-making, both of which should come before analysts estimate centrality and other standard social network metrics.

**5. Quantifying Network Topography**

A number of metrics exist to quantify the topographical features of networks. In terms of the heterarchical-hierarchical dimension, degree, closeness and betweenness centralization all offer glimpses into how centralized a network is although we need to be careful how we interpret our results. In general, the larger a centralization index is, the more likely it is that a single actor is very central while the other actors are not (Wasserman & Faust, 1994, pg. 176), so they can be seen as measuring how unequal the distribution of individual actor values are. Thus, we need to interpret the various indices in terms of the types of centrality estimated.

An alternative measure recommended by Hoivik and Gleditsch (1975) and Coleman (1964) is the variance of degree centrality found in a network (Wasserman & Faust, 1994, pg. 177, 180-181). Finally, if we are working with directed data, then Krackhardt's (1994) graph theoretical measures of hierarchy can be quite informative. To date, however, most social network analyses of terrorist networks have collected undirected data.

Network density is probably the most commonly used metric tapping into the provincial-cosmopolitan dimension. Unfortunately, network density tends to decrease as social networks get larger because the number of possible lines increases rapidly with the number of actors whereas the number of relations which each actor can maintain is generally limited. Consequently, it is of limited use as a measure. We can use it to compare networks of the same size, but that is about all. An alternative suggested by Scott (2000, pg. 75-76) and de Nooy et al (2005, pg. 63) is to calculate a network's average degree centrality. While it is positively associated with "provincialness" of networks, it is not sensitive to network size, which allows analysts to use it to compare networks of different size.

The small world statistic developed by Uzzi and Spiro (2005) to measure the small-worldness of networks of Broadway musical teams also taps into the provincial-cosmopolitan dimension and is worth exploring in some detail here (Humphries and Gurney (2008) developed the identical statistic apparently independently of Uzzi and Spiro). As noted above, small world networks are those where actors cluster into tight-knit groups and the average path length between them is low (Watts & Strogatz, 1998). Local clustering (*CC*) is measured by taking the average of the proportion of an actor's neighbors who also have ties with one another (also known as ego-network density), while average path length (*APL*) is calculated by taking the average of all the shortest path lengths (i.e., geodesics) between all actors in the network. These measures are then typically normalized by calculating the ratio between them and the *CC* and *APL* of a random network of the same size and density:

$$CC_{Ratio} = \frac{CC_{Actual}}{CC_{Random}} \quad (1)$$

$$PL_{Ratio} = \frac{APL_{Actual}}{APL_{Random}} \quad (2)$$

Thus, the more that a network's $CC_{Ratio}$ exceeds 1.0 and the closer its $APL_{Ratio}$ approaches 1.0, the more it resembles a small world network. Uzzi & Spiro and Humphries & Gurney quantified the relationship between the $CC_{Ratio}$ and $APL_{Ratio}$ by calculating a ratio of ratios, so to speak (what Uzzi and Spiro called "small world *Q*):

$$Q = \frac{CC_{Ratio}}{PL_{Ratio}} \quad (3)$$

Later analysis by Uzzi (2008) found that it was unnecessary to compute small world Q in order to predict the probability that a musical would be a critical and financial success. Instead, all that was needed was the $CC_{Ratio}$. Why? Because the $APL_{Ratio}$ almost always approximated 1.0 and recent research (Everton & Lieberman, 2009) demonstrates that Uzzi's discovery was not an exception but the rule. In most networks the $PL_{Ratio}$ will approximate 1.00. Moreover, because a near-perfect correlation exists between the density of an actual network and the *CC* of a comparable random network, there is no need to generate the latter. Small world *Q* can be estimated by simply calculating the ratio of a network's *CC* to its density.

Unfortunately, at this point we do not know what constitutes a cosmopolitan, provincial, hierarchical and/or heterarchical terrorist network. Table 1 lists the relevant social network measures of the "trust," "operational" and "combined" networks of the Noordin Top Terrorist network, but how these measures compare to those generated by other studies is unclear because there is not sufficient data to make such comparison.

**Table 1**. Comparison of Noordin Top terrorist network's topographical metrics

| Metric | Trust Network | Operational Network | Combined Network |
|---|---|---|---|
| Density | 0.081 | 0.362 | 0.378 |
| Average Degree | 6.557 | 28.228 | 29.443 |
| Clustering Coefficient | 0.356 | 0.751 | 0.763 |
| Small World *Q* | 4.238 | 2.074 | 2.019 |
| Degree Centralization | 21.63% (unconnected) | 41.79% | 40.19% |
| Closeness Centralization | | 42.37% | 40.56% |
| Betweeness Centralization | 17.88% | 6.66% | 5.76% |
| Degree Centrality Variance | 53.487 | 222.556 | 225.639 |

**Noordin Top's Terrorist Trust Network**

The Noordin Top Terrorist Trust Network data are drawn from the International Crisis Group's (2006) report on the terrorist networks of Noordin Mohammed Top, who is believed to be responsible for the 2003 JW Marriott Hotel and 2004 Australian Embassy bombings in Jakarta, the 2005 Bali bombing and the 2009 JW Marriott and Ritz Carlton bombings in Jakarta. The initial data were collected and coded by students as part of the "Tracking and Disrupting Dark Networks" course offered at the Naval Postgraduate School in Monterey, California, under the supervision of Dr. Nancy Roberts. Portions of the data have been updated by students in subsequent iterations of the course (through the Spring of 2009) as well as from other articles and reports by Dr. Sean Everton. One and two-mode network data were collected on a variety of relations (e.g., friendship, kinship, internal communications) and affiliations (e.g., schools, religious, businesses, training events, operations). I constructed three one-mode, multi-relational networks (trust, operational and combined) based on the relations listed below. Dichotomized versions of the networks were used to calculate metrics:

*Trust-Network*

- Friendship: Defined as close attachments through affection or esteem between two people. Friendship ties do not include ties based on meetings and/ or school ties.

- Kinship: Defined as a family connection based on marriage. It includes current marriages and past marriages due to divorces and or deaths.

- Religious Affiliation: Defined as association with a mosque. It does not include Islamic schools – see next category – even though such schools have mosques.

- School Affiliation: Educational relations are defined as schools where individuals received formal education. This includes both religious and secular institutions.

*Operational Network*

- Internal communications: Defined as ties based on the relaying of messages between individuals and/or groups inside the nework through some sort of medium.

- Logistical place (Defined as key places where logistical activity – providing materials, weapons, transportation and safehouses occurred.

- Operations: Includes terrorists who were directly involved with the Australian Embassy bombing, the Bali Bombing, the Bali II bombing and/or the Marriott Hotel bombing, either at the scene (e.g., suicide bombers, commanders) or as a direct support to those at the scene (e.g., driver or lookout). It does not include ties formed through communications, logistics, or organizations related to the operations.

- Terrorist Financing: Defined as the for-profit and not-for-profit businesses and foundations that employ members of the network.

- Terrorist Organizational Membership: Defined as an administrative and functional system, whose primary common goal is the operational conduct of terrorist/insurgent activities, consisting of willingly affiliated claimant members. Factions and offshoots are considered separate from their parent organization.

- Training: Defined as participation in any specifically designated activity that teaches the knowledge, skills, and competencies of terrorism. It does not include participation in a terrorist sponsored act or mujahedeen activity in places such as Afghanstan, Bosnia, Chechnya or Iraq unless the individuals' presence was to par ticipate in a specifically designated training camp or base in one of these areas.

**6. Conclusions**

In this paper I have argued that while social network analysis has improved our understanding of how terrorist networks organize, it has generally failed to take into account a network's overall topography before crafting strategies for their disruption. Moreover, research to date has tended to focus on identifying actors who either score high in terms of centrality or are structurally located in such a way that they are in a position to broker the transmission of resources through the network. As I have shown, however, available evidence suggests that networks that are too provincial, too cosmopolitan, too hierarchical and/or too heterarchical tend not to perform as well as networks that maintain a balance between these extremes. If these same dynamics hold true for terrorist networks as well, then identifying key actorswithin a network may not always be the best approach.

Clearly what are most needed in the immediate future are additional studies that explore terrorist networks in all their complexity, not only identifying their central actors but also delineating their topographical characteristics. Moreover, future research should include the development of metrics that adequately quantify the effectiveness of terrorist networks. Number and size of attacks are certainly one type of metric to consider; network resiliency is probably another (Milward & Raab, 2008). For our purposes here, however, identifying the most appropriate metric is of less concern than recognizing that only when one or more are identified will we be able to empirically confirm (or disconfirm) the hypothesized relationships theorized in this paper.

**References**

Arquilla, J., & Ronfeldt, D. (2001). The Advent of Netwar (Revisted). In J. Arquilla & D. Ronfeldt (Eds.), Networks and Netwars (pp. 1-25). Santa Monica, CA: RAND.

Brafman, O., & Beckstrom, R. A. (2006). The Starfish and the Spider: The Unstoppable Power of Leaderless Organizations. New York: Portfolio.

Burns, T., & Stalker, G. M. (1961). The Management of Innovation. London: Tavistock.

Castells, M. (1996). The Information Age: Economy, Society and Culture, Vol. I: The Rise of the Network Society. Malden, MA: Blackwell Publishers.

Coleman, J. S. (1964). Introduction to Mathematical Sociology. New York: Free Press.

de Nooy, W., Mrvar, A., & Batagelj, V. (2005). Exploratory Social Network Analysis with Pajek. Cambridge, UK: Cabridge University Press.

Everton, S. F., & Lieberman, S. (2009). Parsimony and the Quantification of Small World Networks. Unpublished Paper. Monterey, CA: Naval Postgraduate School

Finke, R., & Stark, R. (2005). The Churching of America, 1776-2005: Winners and Losers in Our Religious Economy (2nd ed.). New Brunswick, NJ: Rutgers University Press.

Granovetter, M. (1973). The Strength of Weak Ties. American *Journal of Sociology, 73(6)*, 1360-1380.

Granovetter, M. (1974). Getting a Job. Cambridge, MA: Harvard University Press.

Granovetter, M. (1983). The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory, 1*, 201-233.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology, 91*, 481-510.

Granovetter, M. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives, 19(1)*, 33-50.

Hoivik, T., & Gleditsch, N. P. (1975). Structural Parameters of Graphs: A Theoretical Investigation. In H. M. Blalock (Ed.), Quantitative Sociology (pp. 203-222). New York: Academic Press.

Humphries, M. D., & Gurney, K. (2008). Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence. *PLoS ONE, 3(4)*, e0002051. doi:0002010.0001371/journal.pone.0002051.

International Crisis Group (2006). Terrorism in Indonesia: Noordin's Networks (No. Asia Report #114). Brussels, Belgium: International Crisis Group.

Krackhardt, D. (1992). The Strength of Strong Ties: The Importance of Philos in Organizations. In N. Nohria & R. G. Eccles (Eds.), Networks and Organizations: Structure, Form and Action (pp. 216-239). Boston: Harvard University Press.

Krackhardt, D. (1994). Graph Theoretical Dimensions of Informal Organizations. In K. Carley & M. J. Prietula (Eds.), Computational Organization Theory (pp. 89-111). Hillsdale, NJ: L. Erlbaum Associates.

Milward, H. B., & Raab, J. (2006). Dark Networks as Organizational Problems: Elements of a Theory. *International Public Management Journal, 9(3)*, 333-360.

Milward, H. B., & Raab, J. (2008). The Resilience of Dark Net works in Modern Protectorates – The Case of Liberia. Paper presented at the Workshop on Modern Protectorates, Culture, and Dark Networks, September 19-20.

Nohria, N., & Eccles, R. G. (Eds.). (1992). Networks and Organizations: Structure, Form, and Action. Boston: Harvard Business School Press.

Pedahzur, A., & Perliger, A. (2006). The Changing Nature of Suicide Attacks: A Social Network Perspective. *Social Forces, 84(4)*, 1987-2008.

Pescosolido, B. A., & Georgianna, S. (1989). Durkheim, Suicide, and Religion: Toward a Network Theory of Suicide. *American Sociological Review, 54(1)*, 33-48.

Podolny, J. M., & Page, K. L. (1998). Network Forms of Organization Annual Review of Sociology 1998 (Vol. 24, pp. 57-76). Palo Alto: Annual Reviews, Inc.

Powell, W. W. (1985). Hybrid Organizational Arrangements: New Form or Transitional Development. *California Management Review, 30(1)*, 67-87.

Powell, W. W. (1990). Neither Market Nor Hierarchy: Network Forms of Organization. In B. M. Staw & L. L. Cummings (Eds.), Research in Organizational Behavior: An Annual Series of Analytical Essays and Critical Reviews (Vol. 12, pp. 295-336). Greenwhich, CT: JAI Press, Inc.

Powell, W. W., & Smith-Doerr, L. (1994). Networks and Economic Life. In N. J. Smelser & R. Swedberg (Eds.), The Handbook of Economic Sociology (pp. 368-402). Princeton, N.J.: Princeton University Press.R Development Core Team (2009). R: A Language and Environment for Statistical Computing, from http://www.R-project.org.

Raab, J., & Milward, H. B. (2003). Dark Networks as Problems. *Journal of Public Administration Research and Theory, 13(4)*, 413-439.

Roberts, N., & Everton, S.F. (2011). "Strategies for Combating Dark Networks." Journal of Social Structure 12(2). Retrieved from http://www.cmu.edu/joss/content/articles/volume12//RobertsEverton.pdf

Ronfeldt, D., & Arquilla, J. (2001). What Next for Networks and Netwars? In J. Arquilla & D. Ronfeldt (Eds.), Networks and Netwars (pp. 311-361). Santa Monica, CA: RAND.

Sageman, M. (2004a). Statement to the National Commission on Terrorist Attacks Upon the United States, 2006, from http://www.9-11commission.gov/hearings/hearing3/witness_sageman.htm

Sageman, M. (2004b). Understanding Terror Networks. Philadelphia, PA: University of Pennsylvania Press.

Sageman, M. (2008). Leaderless Jihad: Terror Networks in the Twenty-First Century. Philadelphia: University of Pennsylvania Press.

Saxenian, A. (1994). Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, MA: Harvard University Press.

Saxenian, A. (1996). Inside-Out: Regional Networks and Industrial Adaptation in Silicon Valley and Route 128. *Cityscape: A Journal of Policy Development and Research, 2(2)*, 41-60.

Scott, J. (2000). Social Network Analysis: A Handbook (2nd ed.). Thousand Oaks, CA: Sage Publications.

Stark, R. (1987). How New Religions Succeed: A Theoretical Model. In D. G. Bromley & P. E. Hammond (Eds.), The Future of New Religious Movements (pp. 11-29). Macon Georgia: Mercer University Press.

Stark, R. (1996). Why Religious Movements Succeed or Fail: A Revised General Model. *Journal of Contemporary Religion, 11*, 133-146.

Stark, R. (2005). The Rise of Mormonism. Edited by Reid L. Nielson. New York: Columbia University Press.

Stark, R. (2007). Sociology (10th ed.). Belmont, CA: Wadsworth Publishing Company.

Tsvetovat, M., & Carley, K. M. (2005). Structural Knowledge and Success of Anti-Terrorist Activity: The Downside of Structural Equivalence. Journal of Social Structure, 6, No. 2. Retrieved from http://www.cmu.edu/joss/content/articles/vol ume6/TsvetovatCarley/index.html

Tucker, D. (2008). Terrorism, Networks and Strategy: Why the Conventional Wisdom is Wrong. Homeland Security Affairs, 4:2 (June 2008), 1-18. Retrieved from www.hsaj.org

U.S. Army (2007). U.S. Army/Marine Counterinsurgency Field Manual (FM 3-24). Old Saybrook, CT: Konecky & Konecky.

Uzzi, B. (1996). The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review, 61(4)*, 674-698.

Uzzi, B. (2008). A Social Network's Changing Statistical Properties and the Quality of Human Innovation. *Journal of Physics A: Mathematical and Theoretical, 41*, 1-12.

Uzzi, B., & Spiro, J. (2005). Collaboration and Creativity: The Small World Problem. *American Journal of Sociology, 111(2)*, 447-504.

Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge, UK: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective Dynamics of 'Small World' Networks. *Nature, 393*, 409-410.

# Poverty and Sociability in Brazilian Metropolises: Comparing poor people's personal networks in São Paulo and Salvador

**Renata Mirandola Bichir**
*Evaluation Department of the Ministry of Social Development and Fight Against Hunger*

**Eduardo Marques**
*Department of Political Science*
*University of Sao Paulo*

**Abstract**
Urban poverty encompasses multiple dimensions including distinctive patterns of sociability, as we have recently learned from research carried out in the cities of São Paulo and Salvador, Brazil. Starting with preliminary studies focusing on the role personal networks play in the reproduction of urban poverty, this article aims to compare the personal networks of poor people in these two important Brazilian metropolises, focusing on different types of personal network. Preliminary findings reveal a wide variety of of network types, both in São Paulo and Salvador, but also show great similarity between the two cities. Results show that poor people's networks are quite diverse, although in general they are smaller and less diversified in their sociability profiles than middle-class networks. We also confirmed the relevance of the structure of poor people's networks – and their sociability profiles – in explaining social conditions, looking at inclusion in the labor market, income generation and other dimensions (Marques, 2010a).

**Authors**
*Renata Mirandola Bichir* has a Ph.D. in Political Science from the State University of Rio de Janeiro – IESP-UERJ and serves as the Coordinator of the Evaluation Department of the Ministry of Social Development and Fight against Hunger (MDS) while conducting research at the Center for Metropolitan Studies (Centro de Estudos da Metrópole).

*Eduardo Marques* has a Ph.D. in Social Sciences from the State University of Campinas – Unicamp and is Livro-docente Professor of the Department of Political Science of the University of São Paulo (Universidade de São Paulo: USP) as well as researcher at the Center for Metropolitan Studies (Centro de Estudos da Metrópole).

**Notes**
We would like to thank Miranda Zoppi, Maria Encarnación Moya and Graziela Castello for their engagement in the research project "Redes e pobreza", the origin of the data analyzed in this paper.

*Please address correspondence to Renata Mirandola Bichir at renatambichir@yahoo.com.br or Eduardo Marques at ecmarq@uol.com.br*

## 1. Introduction

This article builds on previous research findings about the role social networks play in the reproduction of urban poverty in Brazil, taking into consideration poor individuals' access to goods and services as obtained through market or social support (Marques, 2010). After studying poverty in Brazilian cities through an approach that was more socio-demographic (CEM-CEBRAP & SAS-PMSP, 2004, Marques & Torres, 2005), we designed a research study to test the joint effects of networks and segregation on poverty conditions, first in São Paulo and then in Salvador. In this research study, we analyze relational structures (the networks), their organization (sociability profiles) and their mobilization in everyday life situations.

This article explores the diversity of poor people's personal networks in São Paulo and Salvador - both large and significant Brazilian metropolises, although the first holds national prominence whereas the second is the most important center in the Northeast region. We would expect very different personal network patterns in both cities, based on the large differences between them, their urban structure, labor markets, family structures, migration patterns, economic structure, daily sociability, amongst other factors. However, in line with studies which have shown, within industrialized countries, similar personal network characteristics in different urban contexts (Fischer and Shavitt, 1995; Grossetti, 2007), we found great similarities in the personal networks of poor people in São Paulo and Salvador. As we shall see later, poor people's networks are quite diverse in their structure and sociability patterns, but the typologies that organize this diversity are more or less the same in São Paulo and Salvador.

Excluding this introduction and the concluding remarks, the article is divided into three sections. The next section briefly discusses recent literature on personal networks, stressing their relevance for understanding poverty and the importance of analyzing full networks instead of ego-centered ones. The second section describes case studies and research strategies. The third section explores the variability of poor people's personal networks through two typologies; one based on their personal network structure and the other on their sociability patterns. In the final considerations we summarize the main findings, stressing the relevance of both typologies for understanding social conditions amongst poor people.

## 2. Poverty and Personal Networks

Social network analysis is a relatively recent component of social science, but its relational ontology has been at the heart of social science since the classics, such as the works of Simmel (Emirbayer 1997). More recently, however, the development of social network analysis methods has enabled the production of precise studies about the effects of relational patterns on a wide variety of processes (Freeman 2004). Although some interesting analyses have been published using networks metaphorically (Fawax 2007 and Gonzalez de la Rocha 2001), the full potential of relational ontology is seen in its methodological use. In discussions about living conditions and poverty, in particular, the international literature has increasingly emphasized the role networks play in access to opportunities (Briggs 2005a, 2005b and 2003), in the presence or absence of a sense of belonging (Blokland and Savage 2008) and in mediating the access individuals and groups may have to three other sources of welfare – markets, sociability and the state (Mustered, Murie and Kesteloot 2006).

While we are particularly interested in poverty dynamics in urban contexts, it is also important to consider the role networks may play in combating the pernicious effects of residential segregation. As highlighted by authors from different traditions, such as Nan Lin (Lin, 1999), Loic Wacquant (Wacquant, 2007), Xavier Briggs (Briggs 2005a, 2005b and 2003) and Talja Blokland (Blokland, 2003), the isolation effect of segregation may be counterbalanced by social ties that can bridge spatial separation, which emphasizes the necessity of integrating social networks into segregation studies. The interaction of networks with segregation and poverty involves the incorporation of informal elements recently highlighted in the urban poverty literature (Mingione 1994, Roy 2005 and Pamuk 2000).

This entire research, therefore, begins with the idea of sociability as a central issue in understanding conditions of urban poverty. Although this statement may appear self-evident, the majority of writings about urban poverty, both in Brazil and abroad, have been polarized by perspectives based on systemic economic dynamics on the one hand and an analysis of individual attributes and behavior on the other (Marques, 2012). Although we agree with the importance afforded to economic conditions, the labor market and individual behavior, we nevertheless believe that societal elements and midlevel processes associated with the relational patterns in which individuals are embedded are highly important for an understanding of poverty.

In order to support this point of view, we researched the personal networks of individuals in diverse situations of urban poverty, firstly in São Paulo and then in Salvador, reconstituting their attributes and relational patterns and investigating the conditioners, consequences and mobilization of their personal networks. Contrary to most of the literature on networks and social support, we sought to analyze the total variety of sociability that poor people may develop in their everyday lives and did not restrict the analysis to family or kin ties, as noted by Lonkila (2010). Furthermore, it is important to note that we analyzed entire personal networks, rather than limiting our analysis to ego-centered networks.

The distinction between egonets and personal networks is not canonic in the literature[1] , although it is also established by works such as Lonkilla (2010). Since several important processes in everyday life come from social connections located at more than one step from the egos, we establish here a distinction between ego-centered networks and personal networks. We consider that ego-centered networks or egonets "consist of one actor (ego) and all other actors (alters) with which ego has direct relations, as well as the direct relations among those alters" (Knoke and Yang, 2008). Personal networks, differently, are the complete or full networks of the personal sociability of a certain ego. To gather information

[1] Important textbooks on networks do not distinguish personal networks from egonets. See for example Wasserman and Faust (1994), p. 41 and 42 and Degenne e Forsé (1994), p. 29.

about the later, we adapted full network data collection tools, considering personal sociability as the topic or theme around which the interview questions were constructed. We developed this research design since we think it is important to analyze poor people's network structures in order to analyze several poverty reproduction mechanisms. This methodological procedure differs from the majority of the literature on social support, which departs from survey information. In spite of the relevance of ego-centered networks obtained from survey data for international comparisons, we agree with Lonkilla's (2010, p.47) criticism: "these surveys do not include data on alters' interconnections and therefore do not enable an analysis of personal network structure". In this sense, our findings are not directly comparable with those presented by the majority of the personal networks literature, whose studies are generally based on ego-centered networks collected in survey researches. However, some general trends are comparable.

In broad terms, the research agenda on sociability and networks departs from the seminal contribution of Wellman (1979), approaching the community question. The author analyzed survey data on closest contacts in the upper–working-class neighborhood of East York, Toronto, considered one the most solidary areas of the city. Wellman showed that intimate networks were prevalent and responsible for social support in emergency situations and in everyday life, although these were mainly provided by only some very close contacts. These networks were composed by both kin and non-kin ties, manly local ones. The results were used to discuss the transformations of community considering contemporary sociability.

The author recently continued to pursue this agenda, considering the transformations of daily sociability in recent decades, though departing from Simmel's classical works on urban life. Based on surveys carried out in Canada, Wellman (2001) points out that new means of communication and transport help to overcome the physical barriers defined by neighborhoods or communities. The author argues that communities are not always embedded in neighborhoods, but become increasingly embedded in social networks. The decline of space did not lead to the end of community in general, but to their transformation from door-to-door to place-to-place. In this way, most people have their own "personal community" and obtain information, help and a sense of belonging from those who live in other locations (Wellman, 2001). As we will see in this article, despite Wellman's hypothesis, in poverty areas space still makes a difference, and personal networks are still strongly embedded in local ties, especially associated to family and to neighborhood, in contexts of high homophily.

Fischer and Shavit's (1995) study is another important reference for comparative studies that consider egocentric networks in urban areas[2]. Starting from a survey conducted in Northern California in the late 1970s and a comparative study carried out in the Haifa region in Israel at the beginning of the 1980s, the authors concluded that networks did not differ greatly in both places. The Israelis, however, presented denser, more kin-based and longer-lasting networks. Fischer and Shavit explained these particularities by contrasting American individualism with Israeli group orientation. Personal networks would thus be affected by different cultural backgrounds: "societal structures and cultures can selectively affect particularities of personal life" (Fischer and Shavit, 1995, p. 143).

In order to verify Fischer's argument in another urban context, Grossetti (2007) transposed the Northern California survey to Toulouse (France). He found convergence in network density between Toulousains and Californians, which he explained was due to the stable relational structures in industrialized countries: these relational structures are not highly sensitive to context variation. In another interesting discussion, Grossetti considers the role that "network capital" plays in reinforcing several forms of inequality, since in Toulouse he found differences in network structures – size, density and homophily[3] – according to the group's level of education. When looking at network differences according to social group – individuals who are poor or middle class – we found similar results in our research, which led us to focus on the various relational mechanisms that may foster poverty reproduction (Marques, 2012).

Another very interesting analysis is provided by Lonkila (2010), who examined the role of work-related social ties as a source of social support for Russian and Finnish workers and then compared these results with China. Based on full personal networks obtained from interviews conducted in Helsinki in 2003 and St. Petersburg in 2000, the author, in line with Ruan et al (1997), found that in both China and Russia the socialist past is still visible in the role of the co-worker within support networks. Taking all the indicators into account, co-workers are more important as a source of social support in Russia than in Finland. Lonkila explains the differences between Russia and Finland by looking at the different impact the workplace has on everyday life and the different life trajectories seen in both cases. The author sees a complex combination of Soviet traditions and post-Soviet experiences in the Russian case, stressing, however, that "the economic aspects alone can hardly explain the observed differences" (Lonkila, 2010, p.54). Similarly, as discussed later, the different economic structures of São Paulo and Salvador do not seem to play a major role in shaping poor people's personal networks in either metropolis, since they have quite similar patterns overall.

Correspondingly, Lee, Ruan and Lai (2005) contrasted socialist and capitalist influences on ego-centered networks when they compared the composition of social support networks in Beijing and Hong Kong, two modern urban societies with a similar cultural heritage but very different social structures. The authors stress the different personal support networks in Western societies and in China, showing that kin is much more relevant in China, although it is relevant in the West. Looking at Beijing and Hong Kong, the authors found great similarities that were explained by the common cultural heritage. They also highlighted close kin as the most important source of social support in both cities. An important dif-

[2] They also analyzed rural areas, but we focus here on the results concerning urban sites.

[3] Homophily is the network characteristic that describes the existence of relationships among individuals with similar attributes. For example, a relationship between two women is homophilic in regards to gender and a relationship between two poor individuals is homophilic in terms of social group. For a detailed analysis of the elements associated with this important relational issue see McPherson et al (2001).

ference was based on income, which is a much more relevant dimension in Hong Kong than in Beijing: in Hong Kong those with a lower income are much more likely than their Beijing counterparts to have no one to turn to across all the dimensions of social support. As we shall see, income is a major differential in explaining personal networks in Brazil, since the greatest differences are found between poor people's personal networks and the personal networks of the middle class, regardless of urban context.

As previously noted, when we focus on social support networks, family and kin are very significant dimensions. This is also true in the Brazilian case, as will be outlined later. One interesting study that considers the role of family in personal networks is provided by Bastani (2007). Analyzing middle-class personal networks by gender in post-revolutionary Tehran, Bastani shows that men and women have similar personal networks in terms of size and percentage of kin in the network, but that they differ substantially in gender composition. More educated people have larger networks, as has been seen in other studies and is also true of the Brazilian case. In contrast to other studies, however, older people have larger networks; this is linked to the Iranian family structure and the resilience of children and siblings within networks, whatever their age.

Several of these dynamics are explained by the macro-structural conditions in post-revolutionary Iranian society. According to Bastani, because of "the socio-economic constraints imposed by the wider society, family and kinship are meeting places where socialization of the young and marriage are facilitated" (2007, p.371). Furthermore, as we also see in the case of poor people's social networks in Brazil, kin ties are very important for the allocation of resources and for various kinds of social support throughout the individual's life: "In Iran, the family serves as an economic and political institution as much as a social one, and individuals maintain close ties to their kin throughout their lives" (Bastani, 2007, p. 372). The relevance of kin is also explained by Iran's young population and high fertility rate, which is no longer the case in Brazil. As we shall see, some of these general trends can be seen in Brazil, a developing country whose society is marked by strong inequalities.

## 3. Case Studies and Methodology

According to the 2008 National Household Sample Survey (Pesquisa Nacional por Amostra de Domicílios: PNAD), Brazil has a population of around 190 million people; 83.75% located in urban areas. Since the 1950s, the accelerating urbanization process has been associated with the continuation of high levels of absolute and relative poverty, both in rural and urban areas. Despite recent decreases in poverty and inequality – due to several combined social and macro-economic policies – it is still important to analyze poverty dynamics, especially in large metropolises such as São Paulo and Salvador.

São Paulo is the largest and most important metropolis in both Brazil and Latin America; in 2010 there were approximately 11 million inhabitants in the municipality and 20 million in the metropolitan region. The city of São Paulo is considered the most important financial and corporate hub of Latin America and in 2005 the city was responsible for 12.5% of Brazil's GDP, 36% of the total production of goods and services of the State of São Paulo and was home to 63% of the multinationals established in Brazil.

In São Paulo one can find both a significant part of the most modern productive activities associated with globalized business and a large poor population living in mainly segregated spaces and with precarious access to services and policies; a clear illustration of Brazilian inequality. Salvador, the capital city of the State of Bahia, is a metropolis of almost three million inhabitants and is the most densely inhabited city in the Northeast region, which in turn is the poorest region in the country and the source of most of the poor's intra-national migration. Salvador is the economic center of the state and an exporting harbor, industrial hub and tourism center. As well as social inequalities, the capital of Bahia also suffers from sex tourism; high levels of unemployment and violence; poor access to public services; and a process of urban sprawl.

### *3.1 Participants*

In order to compare poor people's personal networks in the two cases, the study included two fieldwork phases, one in 2006/2007 in the metropolitan region of São Paulo and the other in 2010 in the City of Salvador. In São Paulo, network interviews were conducted with 209 individuals in seven carefully chosen locales, taking previous studies of urban poverty into consideration in order to include the variability of segregation and housing situations within the city. In this sense, we built an intentional sample of places. The fieldwork included downtown slum tenements; favelas on the urban fringe of the metropolis, in very high-income neighborhoods, in middle-class neighborhoods and in an industrial district; a large-scale housing project on the metropolitan fringe; and a fairly peripheral irregular settlement. In Salvador, the fieldworkwas conducted in five locales based on the same criteria, including downtown slum tenements and favelas in two consolidated and two peripheral regions of the city; here the fieldwork examined 153 personal networks. In order to create parameters to compare the networks, we also studied 30 middle-class networks in São Paulo[4]. Table 1 presents the distribution of cases across the several areas in São Paulo and Salvador under study.

In each of these cities, the interviewees were chosen and approached in public spaces and at their homes on weekdays and at weekends. In only a few cases was our presence in these areas mediated by previous researchers who studied the same locales or by local civil associations[5]. After agreeing to be part of the research, interviewees were asked to answer a few questions on their everyday relationships.

---

[4] Middle class was defined in its broad sense, combining income and professional criteria and included liberal professionals, civil servants, those involved in intellectual activities and owners of commercial establishments. The thirty middle-class networks were only used as a parameter and were not fully analyzed.

[5] We are very grateful to Encá Moya, João Marcos de Almeida Lopes, Teresinha Gonzaga, Letizia Vitale, Gabriel Feltran, Rafael Soares and Henri Gerve seau, who helped us indicating interviewees in some of the areas.

**Table 1**. Distribution of cases across several areas in São Paulo and Salvador.

**São Paulo**

| Areas | Number of Cases |
|---|---|
| High-Income Neighborhoods | 30 |
| Slum Tenements (Cortiços da João Teodoro) | 29 |
| Favela on the urban fringe of the metropolis (Vila Nova Esperanca) | 30 |
| Favela in a high-income neighborhood (Paraisópolis) | 31 |
| Favela in a middle-income neighborhood (Vila Nova Jaguaré) | 30 |
| Favela in an industrial district (Guinle) | 30 |
| Peripheral irregular (Jardim Ângela) settlement | 29 |
| Large scale housing project (Cidade Tiradentes) | 30 |

**Salvador**

| Areas | Number of Cases |
|---|---|
| Slum Tenements (Centro Histórico) | 33 |
| Favela in a consolidated area, next to a middleclass neighborhood (Nordeste de Amaralina) | 37 |
| Favela in a consolidated area, next to a mixed neighborhood (Curuzu) | 31 |
| Favela in a paripheral area, very segregated (Bairro de Paz) | 23 |
| Favela in a peripheral area (Novos Alagados) | 29 |

*3.2 Measures*

The interviews collected both relational information and personal network attributes. In each field, basic social attributes such as gender, age and employment status were used to control the sample and avoid bias. Although we did not follow random sampling statistical techniques, a comparison of interviewee characteristics and the population studied did not suggest the presence of bias.

The interviews used a semi-open questionnaire and a name generator. Besides mapping potential rational spheres, such as leisure, work, neighborhood, association, church, etc., the questionnaire covered basic socioeconomic attributes and the individuals' family configuration and migratory and occupational trajectories. Based on this first questionnaire, we start the definition of the main relational spheres each interviewee had in his/her everyday life. For instance, when someone mentioned friends who get together every weekend to play soccer, this was organized afterward as a leisure sphere.Subsequently, a two-step name generator was used. The interviewee was first asked to list up to five people in each of his/her spheres of sociability – family, neighbors, friends, work colleagues, those from religious centers, associations, places of leisure and others that had emerged during the first part of the interview. In this sense, the several social spheres were identified by the researchers based on the narratives the interviewees provided in the first part of the interviews. Departing from the richness of social ties people mentioned in their everyday life, the spheres were classified following analytical and comparative categories defined by the researchers.

The classification of the spheres, therefore, is cognitive, but was later reorganized in a smaller number of categories by the researchers. The sphere of sociability represents the different social spaces in which the interviewee considers that each of his/hers relationships are mobilized. It does not represent the type of tie or a physical place, but a cognitively defined region of the sociability of each ego (Marques, 2012). An ego may consequently consider, for example, that a relative does not belong to the family sphere but to the leisure sphere, if the tie between them is only mobilized during leisure activities.

Conversely, someone may consider that a friend from the neighborhood belongs to the family sphere because the tie between them is mobilized in his/hers relative's house. In concrete terms, the spheres include certain sets of individuals and organizations, the relationships established between them (of various types and constantly changing), as well as identities, sets of signs and discursive patterns, as interpreted by Mische and White (1998) and White (1995). One could say that, as defined here, the spheres include the more stable versions of Mische's netdoms (2007). In some cases, the spheres may overlap through the existence of individuals who participate in more than one sociability context at the same time. As noted in other studies (Lonkila, 2010), this method enables investigation of the totality of the respondents' daily social relations, including friends, relatives, co-workers, etc., mixing different sociability spheres but not taking into account the types of tie between the nodes, which would be extremely difficult to capture, given that we are analyzing full networks.

The first names quoted in each sphere represented the network 'seed' and were included in the first column of the relational questionnaire. The interviewee was then asked to list up to three names for each of the names in the seed which were associated in his/her mind with the name first cited in terms of their sociability. He/she could present a new name, repeat names, include his/her own name or say no one. These people were included in the rows for each cited name, but the new names were also included in the first column, at the end of the list. With the 'seed' names finalized, the interview went on to the recently added names, producing a snow-balling process within the same interview. The procedure was repeated up to four times (including the seed), but none of the poor individuals reached this limit, suggesting that the frontier of the network had been reached. Following this, the interviewee was asked to classify these people according to two attributes: place of residence (local/non-local) and sphere of sociability in which the tie occurred[6].

Thus we arrived at full personal sociability networks, rather than ego-centered networks[7] . After processing the relational data and constructing the networks for each city, we returned to the field to perform qualitative interviews with selected individuals, combining types of networks and personal characteristics to form a subset. We interviewed 17 individuals in São Paulo and 21 in Salvador, exploring network transformations and, principally, network mobilization where social support is accessed to solve daily problems, such as migration, getting jobs, child and elderly care, emotional support, etc. Network mobilization in São Paulo was analyzed in detail by Marques (2010 and 2012).

Table 1 summarizes the main relational information concerning poor people's personal networks in Salvador and São Paulo and middle-class networks in São Paulo. Firstly, we highlight the great differences between the personal networks of middle-class individuals and those of the poor. Middle-class individuals tend to have larger networks in terms of the average number of nodes and ties. These middle-class results are in line with Grossetti (2010), who demonstrated that the more educated people are, the larger their personal networks. Middle-class individuals also have much more diversified sociability than the poor: middle-class networks have, on average, 5.5 sociability spheres as opposed to 3.8 spheres among poor individuals in São Paulo and 3.5 in Salvador. Considering the relative proportion of ties that fall within institutional spheres (such as work, school, church and civic associations), the sociability of middle class individuals is more likely to be centered around these spheres than the sociability of poor individuals. Evidently, the opposite occurs when considering primary sociability spheres (family, neighbors or friends): the ties of poor people in São Paulo and Salvador are concentrated in these spheres at an average of 77.37% and 79.31%, respectively, while among middle-class individuals the average proportion of this type of tie is 54.58%. When looking at the proportion of ties people establish within the neighborhood in which they live, which we call level of localism, it is interesting to note that, in line with other studies (Grossetti, 2007), middle class networks are much less local than those of poor people.

Poor people's networks reveal similar characteristics in São Paulo and in Salvador, especially when compared to middle class networks. However, we found greater localism in Salvador; slightly larger networks with more varied sociability in São Paulo; and higher relational activity in Salvador (networks have, on average, less nodes but more ties). Besides the fact that smaller networks tend, by definition, to have a higher probability of being better connected, due to the smaller number of nodes, these differences between the cities may be caused by the higher localism in Salvador, leading to networks that are at the same time smaller and more intensely connected. To explore this variability, a typology based on both network characteristics and sociability profiles was developed, as we shall see in the next section.

## 4. Types of Networks and Sociability

As indicated in the preliminary findings, poor peoples' personal networks in São Paulo and in Salvador display a great diversity of patterns and a significant variability in terms of size, sociability sphere and localism, amongst other dimensions. As well as describing the main characteristics of these personal networks, it is important to classify them in order to make a more in-depth analysis of their conditions. Two complementary cluster analyses were undertaken in order to classify these networks.

Firstly, they were classified by looking at several network measures currently used in the network analysis literature. Secondly, networks were classified according to their sociability profiles, taking into account the relative distribution of nodes within different spheres of sociability: family, neighbors, friends, work, religiosity, leisure and civil association. While the first typology aims to explore the networks' main structural characteristics, the second provides information about how the various types of ties manifested by poor people are organized and mobilized in everyday life. This section presents network type first followed by sociability type. In the last part of the section, the two typologies are combined in order to explore the different relational settings which are illustrated with actual cases from São Paulo and Salvador.

### *4.1 Network Types*

In order to analyze and classify the heterogeneity of personal networks in the two cities, 362 networks[8] were submitted to a cluster analysis based on a variety of social network analysis measures: number of nodes, number of ties, diameter, average degree, centralization, clustering coefficient, E-I indexes, n-clans, betweenness, information, structural holes, number of contexts

[6] In the case of São Paulo, people were also asked about the context of sociability in which the tie was created, but since this information did not produce interesting results, it was discarded prior to the Salvador fieldwork.

[7] For operational research reasons a limit was placed on the number of rounds of interviews, which theoretically placed limits on the size of the chosen networks. However, in the case of the individuals living in poverty, the name generator reached the networks' edge before this point and, as such, we may consider the constructed networks as approximately corresponding to a representation of the interviewee's whole networks.

[8] 209 cases demonstrated complete relational information in São Paulo and 153 in Salvador.

and number of spheres[9]. The automatic solution of this analysis generated six groups, which were reclassified into five main types of network, considering that one of the groups included only a few very large networks, which were aggregated into the second largest network group. The groups were particularly varied in terms of size – number of nodes and ties.

The average number of spheres drops slightly when we move from large to small networks. The first two network types demonstrate a similar level of localism: about 68%. There is a similar, although slightly higher, level in the third and fifth types: 73%. The fourth type of network, medium to small, presents much less localism and only 46% internal ties. Figure 1 shows these main characteristics.

Table 2 presents the distribution of network type by city. As can be seen in the first rows, the distribution is quite similar in both cities, although São Paulo presents a slightly higher concentration of small networks, while Salvador presents a slightly higher concentration of large ones. The table also indicates that midsized networks tend to be more common, although distribution is skewed towards small networks.The main aspects of each type of network are briefly outlined below.

*4.2 Large Networks – 34 Cases*

This is the least frequent type of network. Large networks are more common among men, non-migrants and single individuals, as well as those who live in segregated areas. Individuals with this type of network tend to have a higher level of education, which is consistent with the higher concentration of students and young people. Employees with formally registered jobs are overrepresented in this type of network, as are individuals who work outside the neighborhood in which they live and people who participate in some kind of civil association. Levels of precariousness are slightly above average in this group, particularly as a result of family and income precariousness – on average individuals classified in this network type have a lower per capita family income[10]. This is consistent with the higher levels of access to the main federal Conditional Cash Transfer (CCT) welfare program, the Family Grant (Bolsa Família)[11], amongst people who have this kind of network.

*4.3 Large to Medium Networks – 69 Cases*

Women are strongly overrepresented in this type of network as are non migrants and people who are single. People with higher – secondary – levels of education are more likely to have this kind of network, although the group's average income is slightly below the overall average. Civil servants, non-formal workers and the unemployed more frequently have this type of network. Family and housing precariousness are more common amongst people with large to medium networks.

*4.4 Medium Networks – 105 Cases*

This is the most common type of network, representing almost one third of all personal networks. People with this kind of network display demographic characteristics – sex, age, schooling, income and migratory status – similar to the overall average. Married people, house wives, small business owners and people who work in the locality in which they live are all overrepresented in this group. Family, work and income precariousness are more common among individuals with medium networks.

*4.5 Medium to Small Networks – 97 Cases*

Medium to small networks are the second most frequent type of network, representing 27% of all personal networks. As with the previous type, individuals with medium to small networks display average demographic characteristics close to the overall average, especially when considering age (37 years old) and schooling (6.4 years of study). However, individuals within this group have the highest income - almost one minimum wage per capita. This type of network is more frequent amongst older migrants, who have been in their place of residence for more than 10 years; married individuals; those who work in family businesses; formally employed workers, including those in domestic service; and precarious self-employed workers, who mainly work outside the community in which they live. Individuals in this type of network present low levels of all types of precariousness except housing.

*4.6 Small Networks – 56 Cases*

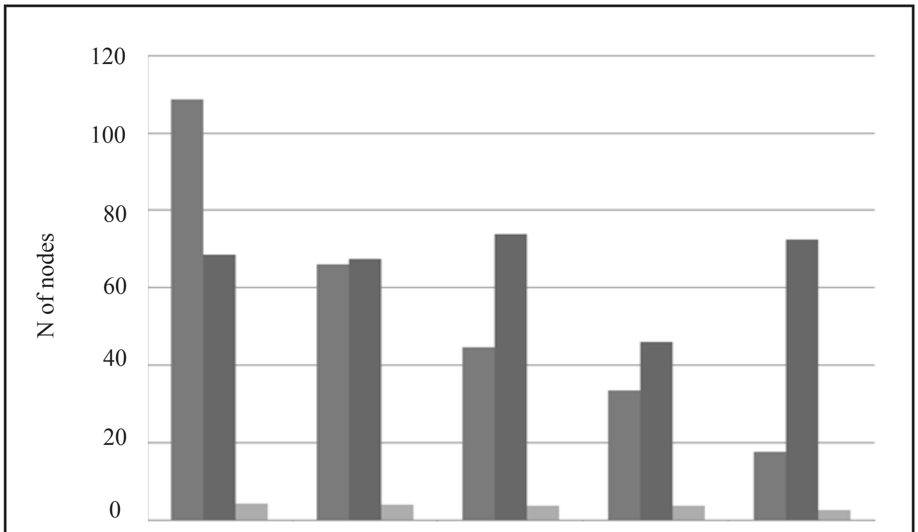
This is the second least frequent type of network, representing 15% of all personal networks. Individuals in this group are, on average, older – 41 years old – and their schooling and income are below the overall aver age. Men, migrants and married people are more likely to have this type of network. Small business owners, retired and unemployed people are also overrepresented in this group, which has a higher concentration of people who work in the locale in which they live. Family, work and income precariousness are more common within this group.

---

[9] All these measures were subjected to cluster analysis using SPSS 13.0 software to apply K-means algorithm. For details on the measures, see Wasserman and Faust (1994).

[10] Income precariousness is present when the average per capita family income is less than or equal to ¼ minimum wage; family precariousness is found where the family nucleus consists of a single adult with small children; housing precariousness is noted when people live in a small shanty house (shack) or, in the case of tenements, in a room without a bathroom; labor precariousness is defined when wages are earned informally, from odd jobs or unregistered unemployment.

[11] The Family Grant (Bolsa Família) Program was created in 2003, during President Lula's first term, as the result of the integration and expansion of several prior CCT programs and is now one of the largest CCTs in the world. In 2010 the program reached 12.6 million families. The program is considered to be one of the causes of the recent drop in Brazil's poverty and inequality levels.




| | Large | Large to Medium | Medium | Medium to Small | Small |
|---|---|---|---|---|---|
| N of nodes | 109 | 66 | 45 | 34 | 18 |
| % of Internal Ties | 68.63 | 67.59 | 73.81 | 45.99 | 72.41 |
| N of spheres | 4.4 | 4.0 | 3.8 | 3.7 | 2.7 |

*Figure 1*. Size, Localism and sociability sphere according to type of network

*4.7 Sociability Type*

As well as classifying personal networks according to their structural characteristics, we clustered them according to the most frequent sociability type, i.e., the prevalence of spheres - family, neighborhood, friendship, church, work and others - was examined within the everyday life of poor people in São Paulo and Salvador.

A cluster analysis of sociability profiles revealed six main types of sociability, depending on whether they were centered on the family, neighborhood, friends, church, work or associations. We consider the first three types – family, neighborhood and friends – to be more primary and potentially homophilic, while the others – church, work and association – tend to be less homophilic and more often based on ties constructed within organizational settings.

Before we present each group in detail, we should point out that in both cities family and neighborhood are the most important spheres for the majority of the poor[12] . Aside from these primary spheres, however, important portions of their sociability are organized in other spheres, as seen in the significance of the six types of sociability presented in Table 3. This presents the distribution of each sociability sphere by type of sociability and highlights any above-average concentrations. The distribution of type of sociability across the two cities (Table 4) once again displays a relatively equal distribution. However, family and friendship-centered networks are more frequent in Salvador, while church, work and association-centered networks are more common in São Paulo.The social situations typically associated with each kind of sociability are described in Table 3.

**Table 2**. Network type by city

| | Large Network | Large to Medium Network | Medium Network | Medium to Small Network | Small Network | Total |
|---|---|---|---|---|---|---|
| **São Paulo** | 8.6% | 18.7% | 27.7% | 30.2% | 14.8% | 100.0% |
| **Salvador** | 10.5% | 19.7% | 30.9% | 22.3% | 16.4% | 100.0% |
| **Total** | 9.4% | 19.1% | 29.1% | 26.9% | 15.5% | 100.0% |
| **# of Cases** | 34 | 69 | 105 | 97 | 56 | 361 |

[12] This is also the case for the family sphere among the middle-classes. See Marques (2012).

**Table 3** . Sociability types by sphere of sociability (%)

| SPHERE | Family | Neighborhood | Friendship | Church | Work | Association | Total |
|---|---|---|---|---|---|---|---|
| **Family** | 64.07 | 28.75 | 37.41 | 33.34 | 31.37 | 34.47 | 40.57 |
| **Neighborhood** | 20.68 | 57.08 | 23.96 | 25.32 | 26.41 | 24.80 | 31.61 |
| **Friendship** | | | 26.22 | | | | 5.89 |
| **Work** | | | | 6.16 | 29.05 | | 8.05 |
| **Leisure** | | | | | | | 1.88 |
| **Church** | | | | 25.02 | | | 4.56 |
| **Association** | | | | | | 19.01 | 1.40 |
| **Studies** | | | | | | | 3.34 |
| **Other** | | | | | | | 1.21 |
| **Number of Cases** | 93 | 86 | 57 | 48 | 55 | 22 | 361 |

**Note**: Percentages below 6% were omitted. Cells highlighted in dark grey have above-average percentages; cells in light grey have important concentrations of some kind of sociability, although below average.

*4.8 Sociability Centered on the Family – 93 Cases*

As previously noted, this is the most common type of sociability: 25% of all personal networks considered in the analysis fell within this type. In fact, there are only 4 poor individuals without any tie within the family sphere; all the others have at least one tie within it. The distribution of this sociability type is fairly even across the two cities and is similar to the average.

When we look at the number of spheres, nodes and ties, the networks of family-centered individuals tend to be smaller than others. The age, schooling and income of individuals with family-centered networks are below the overall average. Women, migrants, married people and illiterate people are overrepresented here, as are housewives, the retired and the unemployed. Catholics and people with no civil participation are more likely to have family-centered networks. Individuals with this pattern of sociability are less exposed to all kinds of precariousness, but have more access to CCT welfare payments than the overall average.

*4.9 Sociability Centered on the Neighborhood – 86 Cases*

This is the second most frequent type of sociability, covering 24% of all poor people's personal networks; only 23 poor individuals – out of the 361 considered – do not have any tie in the neighborhood. There is no difference between São Paulo and Salvador in the distribution of this type of sociability.

Individuals with neighborhood-centered sociability present with average ages, income and schooling levels below the overall average. When we compare them to family-centered individuals, however, we find that neighborhood-centered individuals have higher schooling levels but lower income. Their networks present an average number of spheres; are higher than the average in terms of the numbers of nodes and ties; and present the highest level of localism, as expected. Several demographic characteristics – sex, migratory status – are similar to the average. Single men, precarious self-employed workers, the unemployed and those who work inside their communities are over-represented in this type of sociability.

The same is true of beneficiaries of CCT programs and those who never attend religious centers or civil associations. Neighborhood-centered individuals are more exposed to housing, income and job precariousness; this type of sociability is more frequent in segregated areas.

*4.10 Sociability Centered on Friendship – 57 Cases*

Individuals with friendship-centered sociability represent 16% of all the personal networks of poor people. This type of sociability is seen slightly more frequently in Salvador than São Paulo. Individuals with this pattern of sociability are the youngest and present higher schooling and income levels than the average. Their networks are a little larger than the average, when taking into consideration the number of spheres, nodes and ties. Women, non migrants and single people are overrepresented in this type

**Table 4**. Sociability type by city

| City | Family | Friendship | Neighborhood | Church | Work | Association | Total |
|---|---|---|---|---|---|---|---|
| São Paulo | 25,36 | 23,92 | 14,83 | 13,88 | 15,31 | 6,70 | 100,00 |
| Salvador | 26,32 | 23,68 | 17,11 | 12,50 | 15,13 | 5,26 | 100,00 |
| Total | 25,76 | 23,82 | 15,79 | 13,30 | 15,24 | 6,09 | 100,00 |

of sociability, as are students, housewives, public employees and those who work in the neighborhood in which they live. Individuals with this pattern of sociability are less exposed to all kinds of precariousness and tend to live in non-segregated neighborhoods.

*4.11 Sociability Centered on the Church – 48 Cases*

Sociability centered on any kind of religious congregation represents 13% of all cases. It is important to note that it is quite common in Brazil to profess some form of religion, even though many individuals hardly ever – or never – attend any kind of religious service. The type of sociability seen here, therefore, includes people who, as well as professing a religion, have an active involvement in religious activities and ties with people who have the same religion and/or attend the same religious services. This type of sociability is more frequent in São Paulo than Salvador.

Individuals with this pattern of sociability have age, schooling and income levels similar to the average, but their networks are larger than the average when we consider the number of spheres, nodes and ties. Women, older migrants and married people more frequently present this type of sociability. It is also more common among housewives, retired people, people with formal jobs and those who work outside their neighborhood. As expected, evangelicals who worship on a weekly basis are much more common in this type of sociability, as are people who participate in other civil associations. Family precariousness is above average, but all other types of precariousness are below average. This pattern of sociability is more present in segregated areas.

*4.12 Sociability Centered on Work – 55 Cases*

As described in previous sections, most of the poor people in our sample either work – regardless of the level of job protection – or are looking for jobs. A small portion – 15% – of the total, however, actually presents sociability patterns that are rich in work colleagues. The distribution of this pattern of sociability is similar in both cities. As expected, people with work-centered sociability show higher levels of income (the highest) and schooling but average age. Their networks present the lowest level of localism, or few internal ties, a higher number of spheres than the average and a similar number of nodes and ties to the average.

Men, non-migrants and married people are overrepresented in this type of sociability. The same is true for small business owners, those who work in family businesses, formally employed workers, public employees, workers without legal protection and those who work outside their neighborhood. Catholics who do not attend religious services and those who have no participation in civil associations are also overrepresented in this group. Individuals with this pattern of sociability are exposed to very little precariousness of any kind.

*4.13. Sociability Centered on Associations – 22 Cases*

This is the least frequent type of sociability, representing only 6% of all the personal networks of poor people. We have seen in previous sections that few poor people actually participate in any kind of civil association, neighborhood association, political party or other form of association. We now see that having ties within this kind of association is even rarer. This type of sociability is much more frequent in São Paulo than Salvador.

Individuals with this pattern of sociability are of average age, their schooling levels are higher than the average and their income is below average. The number of spheres and nodes are above average, but the number of ties is below average. Men, single people, those who work inside their neighborhood, workers without formal registration, precarious self-employed workers and unemployed individuals are overrepresented in this type of sociability. As expected, those who attend any kind of civil association are strongly overrepresented in this group, but the same is not true when looking at attendance of religious services. Individuals with this pattern of sociability are more exposed to all kinds of precariousness.

**5. Main Relational Situations**

The combination of the two typologies provides interesting information for an analysis of the networks of poor individuals in the two cities. Although there were 30 possible combinations (5x6), only a certain number of combinations were frequently found. We decided to highlight four combinations which resulted in the classification of 92.5% of all personal networks:

*5.1 Primary Sociability within Small Networks (101 cases)*

Case Number 76, from Taboão, São Paulo, illustrates this relational situation. She is a 21 year-old, non migrant, married to someone who was her neighbor. She has completed high school and is now a housewife, with a per capita family income of only ¼ of the minimum wage. Her network has 19 nodes, 21 ties and 3 spheres of sociability: family, neighborhood and friendship. Figure 2 presents this network.

*5.2 Primary Sociability within Medium Networks (72 cases)*

An example of this type of network is seen in Case Number 293, from Novos Alagados, Salvador. She is 37 years old, a native of Salvador and has lived in this segregated neighborhood all her life. She is single, lives with her sister and three nephews and works in her own home as a manicurist. They are also beneficiaries of the Bolsa Família family welfare program, although their per capita family income is 0.4 minimum wages. She is evangelical and attends religious services in her neighborhood every single day. Her network has 43 nodes, 69 ties and 4 spheres: family, friendship, work and church. Figure 3 presents this network.

*5.3 Primary Sociability within Large Networks (63 cases)*

Case Number 75: this 13-year-old girl, born in Bahia but living for the last 2 years in São Paulo (Vila Nova Esperança), is an illustrative São Paulo example. Her parents are still in the Northeast and she lives with her older sister, helping to take care of her sister's young baby. She studies in the same neighborhood in which she lives and has many friends, several of whom are from a Catholic association, although she herself professes no religion.

Her personal network shows 68 nodes, 66 ties and 4 spheres: family, neighborhood, study and church association. Figure 4 presents this network.

*5.4 Institutional Sociability within Medium Networks (98 Cases)*

An illustration of this combination comes from Case Number 366, who lives in the historical area of downtown Salvador. He is a 39-year-old man from Salvador who lives in a tenement in the downtown area, where he also owns a small bar and earns 2.6 minimum wages per capita. His network has 45 nodes, 72 ties and 4 spheres ofsociability: family, neighborhood, work and leisure. Figure 5 presents this network.

We did not find a significant number of cases of institutional sociability – focused on church, work or association – within the small or large networks. Although the first three types – primary sociability with small, medium or large networks – tend to be associated with worse socioeconomic conditions, the last type, institutional mid-size networks, is more commonly associated with better social conditions and attributes. The results are similar if we consider the two cities analyzed separately or polled together.

This result echoes findings from previous studies (Marques 2009a and 2010a) considering only the cases of São Paulo. Those analyses showed, using CHAID techniques and regression models, that certain types of networks appeared to be highly associated with classical elements from poverty studies, such as employment, stable employment, social vulnerability and income, despite traditional variables such as education and household size. The worst social situations were associated with highly homophilic sociability patterns and highly local networks. In contrast, the best social situations were associated with middle-size and non-local networks and with sociability concentrated on organizational spheres (work, church, associations).

Therefore, low homophily and low localism tend to be directly associated with better social situations, although network size did not exert a direct influence. Mid-sized networks, however, tended to be better when combined with less homophilic sociability patterns. It is impossible to determine a strict causality here. Social networks and individual attributes are constructed by biunivocal causality throughout the life trajectories of individuals and are im-

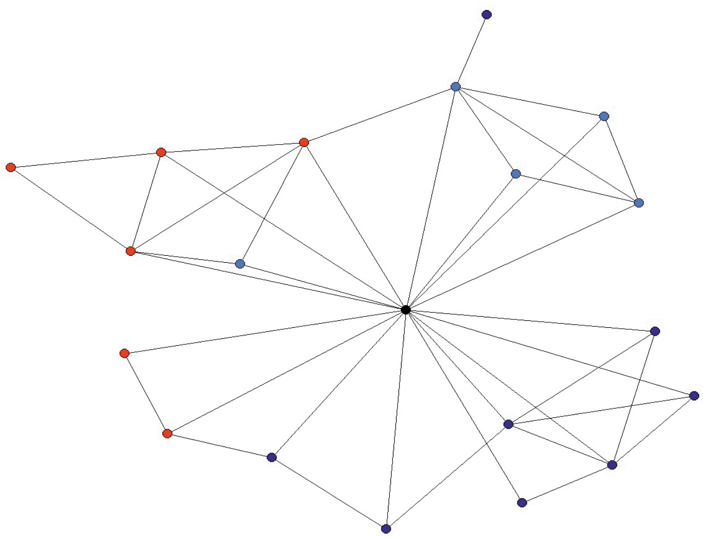


*Figure 2*. Case 76, São Paulo

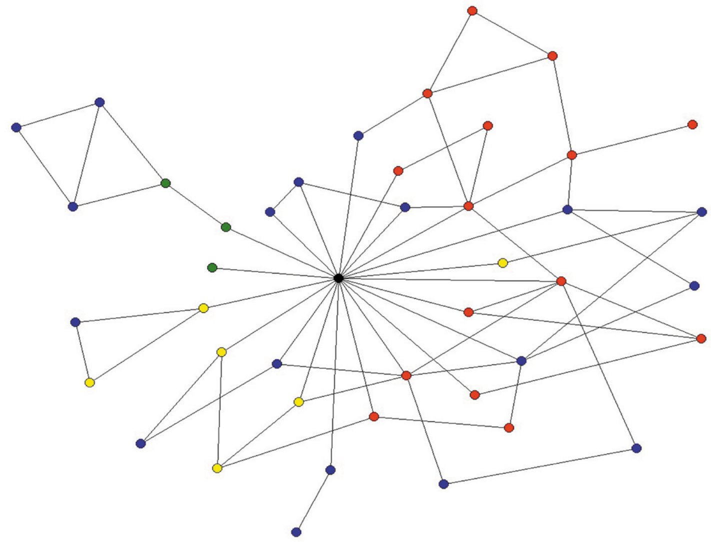


*Figure 3*. Case 293, Salvador

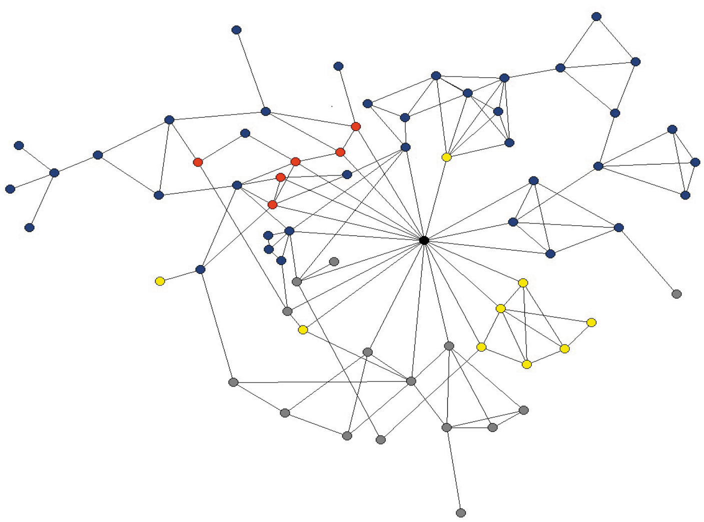


*Figure 4*. Case 75, São Paulo

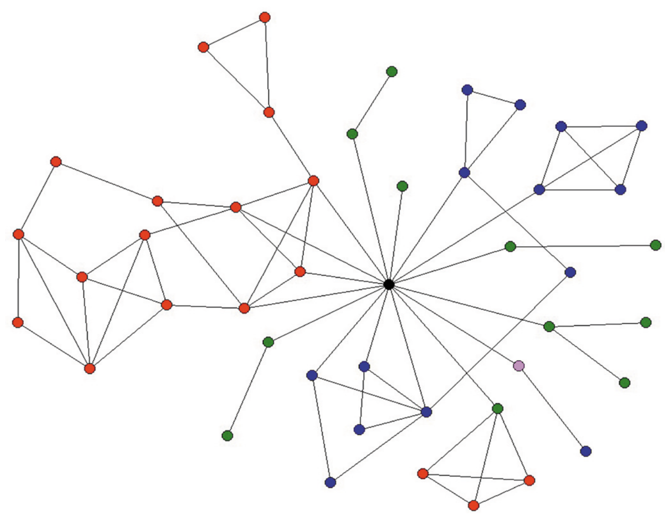


*Figure 5*. Case 366, Salvador

pacted on by individual decisions and events (migration, marriage, divorce, child birth, etc), as well as by the effects of other individuals' networks and decisions.

**6. Final Considerations**

In this paper, we showed that poor people's personal networks tend to present similar characteristics regardless of whether the urban context is São Paulo or Salvador. Despite the differences between these cities, we found very similar patterns in poor people's network size, density and the variability of sociability in both cities. Poor people's networks are more diverse and show greater heterogeneity than one would expect from an economic view of poverty.

The most relevant distinction is found when we compare poor people's network structures with middle-class personal networks. The relational patterns of poor individuals tend to be, on average, smaller, less diverse, more local and more strongly based on primary contacts than middle-class networks. This result suggests that in the Brazilian case there is a metropolitan sociability pattern according to social group, i.e., social class plays a major role in organizing personal networks in Brazil. This great cleavage of sociability patterns by social class and not by urban or cultural context follows previous results of the international literature and strengthens the assumption that social networks reinforce urban poverty reproduction.

When we zoom in, however, we find a great diversity of networks even among the poorest inhabitants of these two important Brazilian metropolises, considering network structures and sociability patterns. As we saw, this diversity may be organized into two typologies, leading to the establishment of five types of networks and six sociability profiles, which are very similar in both cities. The presence of regularities is so relevant that only four combinations of networks and sociability patterns explained the large majority of the cases. In this sense, poor people's networks in São Paulo and Salvador may be organized according to very similar typologies, suggesting a similar urban sociability pattern in both cities, when the focus is on poor people's personal networks. These types are relevant since they are related to different social opportunities, as discussed in Marques (2009a and 2010a). In fact, the best social conditions are associated systematically with middle sized, less local networks and with sociability constructed within organizational settings.

The results are consistent regardless of differences between the two cities, suggesting that they represent a pattern: more local and more homophilic networks are associated with worse social conditions. Since these networks mediate individual access to several kinds of opportunity (in the labor market, daily help and welfare in its broadest sense), these patterns have an important circular characteristic. These circularities reinforce the importance of analyzing poor people's social networks in order to understand poverty reproduction mechanisms and to move beyond the attribute-centered studies and macro-sociological analyses so commonly found in the Latin American urban poverty debate.

**References**

Bastani, S. (2007). Family comes first: Men's and women's personal networks in Tehran. *Social Networks: An international journal of structural analysis, 29(3)*, 357-374.

Blokland, T. (2003). Urban bonds. London, England: Basil Blackwell.

Blokland, T., & Savage, M. (2008). Networked urbanism: social capital in the city. London, England: Ashgate Ed.

Briggs, X. (2003). Bridging networks, social capital and racial segregation in America. HKS Faculty Research Working Paper Series, vol 2, ed. 9, 2-11.

Briggs, X (2005a). Social capital and segregation in the United States. In: D. P. Varady (Ed.). Desegregating the city: Ghettos, Enclaves, And Inequality. Albany, United States: Suny Press.

Briggs, X. (Ed.) (2005b). The Geography of opportunity Race and Housing Choice in Metropolitan America. Washington, United States: Brookings Institution Press.

CEM-CEBRAP & SAS-PMSP (2004). Mapa da vulnerabilidade social da população da cidade de São Paulo. São Paulo, Brazil: Sesc-SP.

Degenne, A., & Forsé, M (1994). Les réseaux sociaux. Paris, France: Armand Colin.

Emirbayer, M. (1997). Manifesto for a relational sociology. *American Journal of Sociology, 103(2)*, 281-317.

Fawax, M. (2008). An unusual clique of city-makers: social networks in the production of a neighborhood in Beirut (1950-75). *International Journal of Urban and Regional Research, 32(3)*, 565-585. doi:10.1111/j.1468-2427.2008.00812.x

Fischer, C. S., & Shavit, Y. (1995). National differences in network density: Israel and the United States. *Social Networks: An international journal of structural analysis, 17(2)*, 129-145. doi: 10.1016/0378-8733(94)00251-5

Freeman, L. (2004). The development of social network analysis. Vancouver, Canada: Empirical Press.

Gonzalez de la Rocha, M (2001). From the resources of poverty to the poverty of resources? The erosion of a survival model. *Latin American Perspectives, 28(4)*, 72-1000.

Grossetti, M. (2007). Are French networks different? Social Networks. *An international journal of structural analysis, 29(3)*, 391-404.

Knoke, D., & Yang, S. (2008). Social network analysis (Quantitative Applications in the Social Sciences; 154). Thousand Oaks, United States: Sage Publications, 2nd ed.

Lin, N. (1999). Building a network theory of social capital. Connections, Vol 22(1), 28-51.

Lee, R., Ruan, D., & Lai, G. (2005). Social structure and support networks in Beijing and Hong Kong. Social Networks. *An international journal of structural analysis, 27(3)*, 249-274. http://dx.doi.org/10.1016/j.socnet.2005.04.001

Lonkila, M. (2010). The importance of work-related social ties in post-soviet Russia: the role of co-workers in the personal networks in St. Petersburg and Helsinki. *Connections, 30(1)*, 46-56.

Marques, E. (2012). Opportunities and deprivation in the urban south: poverty, segregation and social networks in São Paulo. London, England: Ashgate Publishing.

Marques, E. (2010) ¿Como son las redes de los individuos en situación de pobreza en el Brasil urbano? *REDES – Revista hispana para el análisis de redes sociales, 18(9)*.

Marques, E. (2009a). As redes sociais importam para a pobreza urbana? *DADOS – Revista de Ciências Sociais, 52(2)*, 471-505.

Marques, E. (2009b). As redes importam para o acesso a bens e serviços obtidos fora de mercados? *Revista Brasileira de Ciências Sociais, 24(71)*, 25-40. http://dx.doi.org/10.1590/S0102-69092009000300003

Marques, E., & Bichir, R. (2011). Redes de apoio social no Rio de Janeiro e em São Paulo. Novos Estudos CEBRAP, Vol 90. http://dx.doi.org/10.1590/S0101-33002011000200006

Marques, E., & Torres, H. (Eds.) (2005). São Paulo: segregação, pobreza e desigualdade sociais. São Paulo, Brazil: Editora Senac.

McPherson, M., Smith-Lovin, L. & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology, 27*, 415-444. DOI: 10.1146/annurev.soc.27.1.415

Mingione, E. (1994). Life strategies and social economies in the postfordist age. International Journal of Urban and *Regional Research, 18(1)*, 24-45. DOI: 10.1111/j.1468-2427.1994.tb00249.x

Mische, A. (2007). Partisan Politics: Communication and contention across Brazilian youth activist networks. Princeton, United States: Princeton University Press.

Mische, A., & White, H. (1998). Between Conversation and Situation: Public Switching Dynamics Across Network Domains. Social Research. *An International Quarterly of the Social Sciences, 65(3)*.

Mustered, S., Murie, A., & Kestellot, C. (2006). Neighborhoods of poverty: urban social exclusion and integration in Europe. London, England: Palgrave Ed.

Pamuk, A. (2000). Informal institutional arrangements in credit, land markets and infrastructure delivery in Trinidad. *International Journal of Urban and Regional Research, 24(2)*, 379-496. DOI: 10.1111/1468-2427.00253

Roy, A. (2005). Urban informality: towards an epistemology of planning. *Journal of the American Planning Association, 71(2)*, 147-158. DOI:10.1080/01944360508976689

Ruan, D., Freeman, L., Dai, X., Pan, Y., & Zhang, W. (1997). On the changing structure of social networks in urban China. *Social Networks: An international journal of structural analysis, 19*, 75-89.

Wacquant, L. (2007). Urban Outcasts: a comparative sociology of advanced marginality. Cambridge, United States: Polity Press.

Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge, United States: Cambridge University Press.

White, H. (1995). Network switchings and bayesian forks: reconstructing the social and behavioral sciences”. *Social Research: An international quarterly of the social sciences, 62(4)*, 1035-1063.

# Assessing A Novel Approach To Identifying Optimal Threshold Levels For Cognitive Consensus Structures: Implications and general applications

**Jarle Aarstad**
*Faculty of Engineering, Bergen University College*
*Bergen, Norway*

**Abstract**
Previous research has demonstrated the importance of cognitive social network structures to better understanding human behavior and thought. Yet network members may deviate in perceiving whether relations exist between pairs of nodes in a network, which can present a challenge in modeling cognitive consensus structures. It has been suggested to define cognitive consensus structures (CCS) to yield a minimum threshold level of 50% of network members perceiving that a relation exits. Here I suggest an improved operational definition, labeled optimal cognitive consensus structures (OCCS). The OCCS threshold level is a function of a consensus structure; yielding the maximum correlate with the summation of nodes' cognitive interpretations of a social network. Revisiting two datasets, I find that the OCCS' predictive validity outperforms the CCS concept in most cases. I also argue how the OCCS can be further developed as a general tool for optimally dichotomizing valued relational data.

**Author**
*Jarle Aarstad, Ph.D.* is an Associate Professor in Organization Studies at Bergen University College, Faculty of Engineering. His major research interests are social network theory, innovation studies, and entrepreneurship.

**Notes**
This research was supported by grants provided by Bergen University College. Dr. Aarstad would like to thank David Krackhardt for his permission to use the data.

*Please address correspondence to Jarle Aarstad, Faculty of Engineering, Bergen University College, PO Box 7030, Nygårdsgaten 112, NO- 5020 Bergen, Norway. Email: jarle.aarstad@hib.no*

## 1. Introduction

Numerous studies address the importance of cognitive social network structures to better understanding human behavior and human thought (e.g., Kilduff, Crossland, Tsai, & Krackhardt, 2008; Kilduff & Krackhardt, 1994; Killworth, McCarthy, Bernard, & House, 2006; Krackhardt, 1987a, 1990). Cognitive structures are also applied as a proxy for missing data in social network research (Burt & Ronchi, 1994; Krackhardt, 1990; Neal, 2008). But how do we define and decide whether cognitive relations exist between pairs of nodes in a network? Krackhardt (1987a, p. 118) suggests that “a relation exists from i to j if and only if the majority [50% or more] of the members of the network perceives that it exists.” He labels the concept as cognitive consensus structures (CSS).

A threshold value of 50% may be intuitive and make sense in order to model consensus structures, but I argue that such a threshold value can result in suboptimal measures and mask important aspects of the underlying cognitive structure. I therefore suggest an improved operational definition which I label optimal cognitive consensus structures (OCCS). The OCCS threshold level is a function of a consensus structure yielding the maximum correlate with the summation of nodes’ cognitive interpretations of a social network. I will further elaborate this issue below.

I study OCCS on cognitive friendship and advice networks from two classical datasets. I also compare the concept’s predictive validity with Krackhardt’s (1987a) operational definition of CCS. I find that the OCCS’ predictive validity outperforms the CCS concept in most cases. Finally, I argue how the operational definition of OCCS can be further developed as a general tool to optimally dichotomizing valued relational data.

Let us assume that actors a, b, and c are members of a larger network and that there are possible relations between them (for simplicity we now treat the ties as symmetric). Let us further assume that 48% of the network members perceive that a relation exists between a and b, 49% between a and c, and 50% between b and c. Applying a threshold value of 50% would only report cognitive consensus on the b-c dyad. Thus, two out of three relations would be omitted despite the fact that the difference in members perceiving them is marginal. But is a threshold value set to 50% defendable? If not, where should it be set to construct a cognitive consensus structure? Should only the b-c relation be reported? Should we lower the threshold value to 49 % to accept the a-c relation, or even further and also accept the a-b relation?

The best way to deal with this issue, I argue, is to first summarize the network members’ cognitive “slices” or maps of perceived ties into one matrix (or network), which I label the sigma cognitive structure (ΣCS). This aggregated matrix represents the totality of the cognitive interpretation of a network forming the baseline for generating an optimal cognitive consensus structure. Next, we correlate the ΣCS matrix with cognitive consensus matrices yielding different threshold values. I argue that the threshold value generating the highest correlation – i.e., the best fit – embodies an optimal cognitive consensus structure (OCCS). The higher (lower) the correlate, the higher (lower) the correspondence between the matrices.

One might object to using “all” cognitive slices (i.e., the ΣCS matrix) as a baseline for modeling the OCCS. But let us assume that a few nodes’ cognitive structures are 1) totally out of touch with each other and 2) are also totally out of touch with the other nodes’ cognitive maps that, for their part, are totally coherent. In practical terms, the few nodes’ cognitive structures are random and should therefore have a negligible effect on the substantial information that the ΣCS matrix provides. Most nodes’ cognitive maps will lie somewhere between complete randomness and coherency. Accordingly, drawing upon my reasoning, the more random the nodes are in cognitive interpretations, the less weight they will tend to carry in the substantial information that the ΣCS matrix provides. Conversely, the more coherent (i.e., the less random) the nodes are in cognitive interpretations, the more weight they will tend to carry in the substantial information that the ΣCS matrix provides. Accordingly, the advantage of using “all” cognitive slices (i.e., the ΣCS matrix) as a baseline for modeling the OCCS is that it takes account of nodes’ variations in interpretations between randomness and coherency of social network structures.

## 2. Methods and Results

### *2.1 Research Contexts*

The datasets were gathered by David Krackhardt and colleagues from two different firms. The first set was gathered from the managerial group (21 managers) in a 10 year old firm producing high-tech machinery for other enterprises (Krackhardt, 1987a). The data have later been applied in other studies (e.g., Kilduff et al., 2008; Krackhardt, 1987a; Krackhardt & Kilduff, 1999, 2002; Wasserman & Faust, 1994). I denominate the data 21M. The second dataset was gathered from a small entrepreneurial firm, given the pseudonym Silicon Systems. At the time the data was collected the firm had 36 employees and had grown from 3 employees 15 years earlier (Krackhardt, 1990). This dataset has also been applied in numerous studies (e.g., Aarstad, Selart, & Troye, 2011; Bondonio, 1998; Kilduff et al., 2008; Kilduff & Krackhardt, 1994; Krackhardt, 1990; Krackhardt, 1992; Krackhardt & Kilduff, 1999; Wasserman & Faust, 1994). “Silicon Systems’ business involved the sales, installation, and maintenance of the state-of-the-art information systems in client organizations…” (Krackhardt, 1990, p. 347). I denominate the data SiSys.

### *2.2 Data Instruments*

Questionnaires were used to gather data on cognitive advice and friendship ties. Similar procedures were followed for the two populations. I describe the process of gathering data by referring to the SiSys population. In the survey, the section about advice for work-related problems was followed by advice and friendship questions as described by Krackhardt (1990, pg. 349):

> *“...36 questions, (e.g., ‘Who would Cindy Stalwart go to for help or advice at work?’), each asking the same question about a different employee. Each of these 36 questions was followed by a list of 35 names, any of which the respondent could check off in response to the question. Similarly, another section of the questionnaire asked about frienship. The directions for this section paralleled those in the previous section...(e.g., ‘Who would Cindy Stalwart consider to be a personal friend?’), and each question was followed by a list of 35 names, any number of which the respondent could check off.”*

*2.3 Modeling the OCCS and Assessing the Predictive Validity*

In Figure 1 the OCCS in accordance with the previous definition. Regarding the SiSys data, I apply the 33 cognitive "slices" from the participating employees. Figure 1 illustrates different QAP correlates (Krackhardt, 1987b) with ΣCS when modeling the OCCS. For all the cognitive structures we observe that the optimal threshold value is below 50% (0.5), particularly in the case of friendship networks. Said differently, there is less cognitive consensus regarding who is friends with whom as compared to who gives advice to whom.

To be more specific, a threshold value of 23.8% (5 out of 21 employees agree upon a relation) generates the OCCS for the 21M friendship data (correlate .855). A threshold value of 27.3% (9 out of 33) generates the OCCS for the SiSys friendship data (correlate .847). A threshold value of 42.9% (9 out of 21) generates the OCCS for the 21M advice data (correlate .845). And finally, a threshold value of 39.4% (13 out of 33) generates the OCCS for the SisSys advice data (correlate .873).

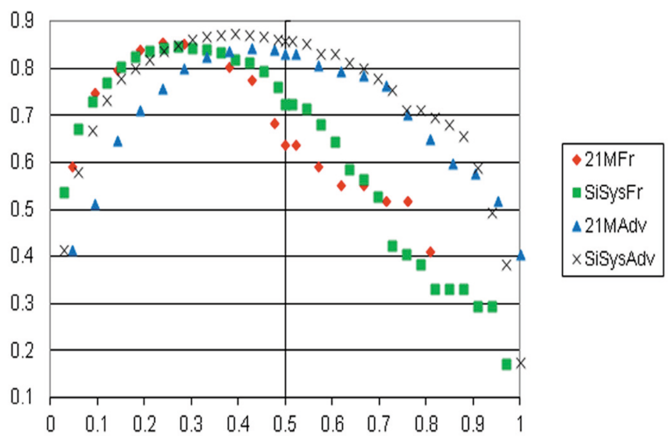



*Figure 1*. Threshold values for modeling OCCS

Predictive validity is the correlation between a given concept and an external criterion (Cronbach & Meehl, 1955). Scholars argue that cognitive and real network structures are expected to be associated (Carley & Krackhardt, 1996; Krackhardt & Kilduff, 2002). To assess the predictive validity of the OCCS, I therefore correlate this structure with the corresponding real network structure; which here represents an external criterion. I model real network ties in accordance with Krackhardt's (1987a) definition of locally aggregated structures (LAS), implying that both *i* and *j* must agree upon a relation between them. In addition, I apply the same criterion to assess the predictive validity of Krackhardt's (1987a) operational definition of cognitive consensus structures (CCS). Finally, I also correlate the LAS structures with the corresponding ΣCS. As noted, for the SiSys data 3 out of 36 employees did not participate in the study, thus I omit these actors – along with possible cognitive and real ties to and from them – when assessing the predictive validity.

Table 1 shows that in three out of four cases, the predictive validity of the OCCS outperforms the predictive validity of the CCS[1]. In particular the OCCS appears to be a preferred measure for friendship data. For the advice data, the results are less clear-cut and somewhat mixed. The findings altogether indicate that when the consensus structure deviates much from a threshold value of 50% (which is the case for the friendship data), the OCCS seems to be a better measure than the CCS. When the consensus structure deviates less (which is the case for the advice data), then by and large either method appears to be equally valid. In addition, we observe that the ΣCS represents a superior measure in terms of predictive validity for all of the four networks (as compared to both the CCS and the OCCS), which indicates that the LAS embody valid external criteria.

**Table 1**. Using LAS structures as criteria to assess predictive validity

| | CCS | OCCS | ΣCS |
|---|---|---|---|
| **21MFr** | .395 | .619 | .675 |
| **SiSysFr** | .453 | .586 | .662 |
| **21MAdv** | .553 | .560 | .637 |
| **SiSysAdv** | .628 | .625 | .680 |

Table 2 reports that the density levels of the CCS are lower than the density levels of the OCCS for all of the networks. This is particularly the case for the friendship data. This is an expected finding, due to the fact that the threshold levels were lower than 50% when modeling the OCCS networks. More importantly, however, we observe that the OCCS and the LAS are fairly similar in network densities whereas the CCS and the LAS are fairly dissimilar. Assuming that the LAS embodies valid external criteria, this indicates that the CCS consistently underreports cognitive relations.

## 3. Discussion and General Application

Studies have addressed the importance of cognitive network structures to better understanding human behavior and thought (e.g., Kilduff et al., 2008; Kilduff & Krackhardt, 1994; Killworth et al., 2006; Krackhardt, 1987a, 1990). Cognitive structures have also been applied as a proxy for missing data in social network research (Burt & Ronchi, 1994; Krackhardt, 1990; Neal, 2008).

**Table 2**. Comparing network densities

| | CCS | OCCS | ΣCS |
|---|---|---|---|
| **21MFr** | .026 | .136 | .121 |
| **SiSysFr** | .028 | .078 | .094 |
| **21MAdv** | .226 | .298 | .307 |
| **SiSysAdv** | .063 | .075 | .094 |

[1]Also here I apply Krackhardt's (1987b) method of QAP correlation, and all reported correlates are strongly significant ($p<.001$; two tailed tests).

Regarding the modeling of cognitive consensus structures, Krackhardt (1987a, p. 118) suggests that "a relation exists from i to j if and only if the majority [50% or more] of the members of the network perceives that it exists." He describes the concept as cognitive consensus structures (CCS). I have argued, however, in this research note that a threshold value of 50% can result in suboptimal measures and mask important aspects of the underlying cognitive structure. Consequently, I have proposed an improved measure, labeled as optimal cognitive consensus structures (OCCS). The OCCS threshold level is a function of a consensus structure yielding the maximum correlate with the summation of nodes' cognitive interpretations of a social network.

Both empirical analyses on cognitive advice and friendship data from two classical cases show that the CCS generates suboptimal cognitive consensus structures, particularly in the case of the friendship networks (Figure 1). Studying network densities, the CCS also consistently appear to underreport cognitive relations (Table 2). Furthermore, for the friendship data, the predictive validity of the OCCS outperforms the CCS. When it comes to the advice data, the predictive validity of the OCCS and the CCS is by and large equal (Table 1). Said differently, the predictive validity of the OCCS outperforms the predictive validity of the CCS in most cases. This is particularly the case when the consensus structure deviates much from a threshold value of 50% (which is the case for the friendship data). All in all, I conclude that the OCCS appears to be the preferred measure over the CCS. To further validate the concept of OCCS, future research should nevertheless perform similar analyses on more types of cognitive data and in different contexts.

On several occasions researchers may need to dichotomize valued relational data, e.g., when studying small worlds (Watts, 1999; Watts & Strogatz, 1998), scale free networks (Barabasi & Albert, 1999), or transitivity in networks (Davis, 1979; Davis & Leinhardt, 1972; Holland & Leinhardt, 1970, 1971). I argue that the approach to modeling OCCS can also be generalized to optimally dichotomizing valued relational data (i.e., replacing valued relational data with either 0 or 1). Let us assume, for instance, that minimum tie strength in a network is $m$ $(0<m)$ and maximum tie strength is $n$. An iterative process can take place to generate different dichotomized matrices yielding cutoff values between $m$ and $n$. The stronger a dichotomized matrix correlates with the original matrix, the better it embodies the original matrix. Thus, the dichotomized matrix yielding the highest correlate with the original matrix defines the optimally dichotomized matrix.

**References**


Aarstad, J., Selart, M., & Troye, S. V. (2011). Advice seeking networks and the learning oragnization. *Problems and Perspectives in Management, 9(2)*, 63-70.

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science, 286(5439)*, 509-512.

Bondonio, D. (1998). Predictors of accuracy in perceiving informal social networks. *Social Networks, 20*, 301-330.

Burt, R. S., & Ronchi, D. (1994). Measuring a large network quickly. *Social Networks, 16*, 91-135.

Carley, K., & Krackhardt, D. (1996). Cognitive inconsistencies and non-symmetric friendship. *Social Networks, 18(1)*, 1-27.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity for psychological tests. *Psychological Bulletin, 52*, 281-302.

Davis, J. A. (1979). The Davis/Holland/Leinhardt studies: An overview. In P. W. Holland & S. Leinhardt (Eds.), Perspectives on Social Network Research (pp. 51-62). New York: Academic Press.

Davis, J. A., & Leinhardt, S. (1972). The structure of positive interpersonal relations in small groups. In J. Berger (Ed.), Sociological Theories in Progress. (Vol. 2, pp. 218-251). Boston: Houghton Mifflin.

Holland, P. W., & Leinhardt, S. (1970). A method for detecting structure in sociometric data. *American Journal of Sociology, 70*, 492-513.

Holland, P. W., & Leinhardt, S. (1971). Transitivity in structural models of small groups. *Comparative Group Studies, 2(2)*, 107-124.

Kilduff, M., Crossland, C., Tsai, W., & Krackhardt, D. (2008). Organizational network perceptions versus reality: A small world after all? *Organizational Behavior and Human Decision Processes, 107(1)*, 15-28.

Kilduff, M., & Krackhardt, D. (1994). Bringing the individual back in: A structural-analysis of the internal market for reputation in organizations. *Academy of Management Journal, 37(1)*, 87-108.

Killworth, P. D., McCarthy, C., Bernard, H. R., & House, M. (2006). The accuracy of small world chains in social networks. *Social Networks, 28*, 85-96.

Krackhardt, D. (1987a). Cognitive social structures. *Social Networks, 9*, 109-134.

Krackhardt, D. (1987b). QAP partialling as a test of spuriousness. *Social Networks, 9(2)*, 171-186.

Krackhardt, D. (1990). Assessing the political landscape: Structure, cognition, and power in organizations. *Administrative Science Quarterly, 35*, 342-369.

Krackhardt, D. (1992). The strength of strong ties: The importance of philos in organizations. In N. Nohria & R. G. Eccles (Eds.), Networks and organization. (pp. 216-239). Boston, MA: Harvard Business School Press.

Krackhardt, D., & Kilduff, M. (1999). Whether close or far: Social distance effects on perceived balance in friendship networks. *Journal of Personality and Social Psychology, 76(5)*, 770-782.

Krackhardt, D., & Kilduff, M. (2002). Structure, culture and Simmelian ties in entrepreneurial firms. *Social Networks, 24(3)*, 279-290.

Neal, J. W. (2008). "Kracking" the missing data problem: Applying Krackhardt's cognitive social structures to school-based social networks. *Sociology of Education, 81(2)*, 140-162.

Wasserman, S. S., & Faust, K. (1994). Social Network Analysis: Methods and Applications (Vol. 8). Cambridge, NY: Cambridge University Press.

Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology, 105(2)*, 493-527.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 394(6684)*, 440-442.

# An Introduction to Personal Network Analysis and Tie Churn Statistics Using E-NET

**Daniel S. Halgin**
*LINKS Center for Social Network Analysis*
*Gatton College of Business & Economics*
*University of Kentucky*

**Stephen P. Borgatti**
*LINKS Center for Social Network Analysis*
*Gatton College of Business & Economics*
*University of Kentucky*

**Abstract**
In this article we review foundational aspects of personal network analysis (also called ego network analysis) and introduce E-NET (Borgatti 2006), a computer program designed specifically for personal network analysis. We present the basic steps for personal network data collection and use E-NET to review key measures of personal network analysis such as size, composition and structure. We close by introducing longitudinal measures of personal network change, including tie churn, brokerage elasticity, and triad change. We argue that these measures can help reveal change patterns consistent with tie formation strategies that would otherwise be missed using more traditional analytic approaches.

**Authors**
*Daniel S. Halgin* is Assistant Professor of Management at the University of Kentucky and a member of the Links Center for Social Network Analysis. He received his Ph.D. in Management from Boston College. His research interests include social network theory and methodology, identity dynamics, and career trajectories.

*Stephen P. Borgatti* holds the Paul Chellgren Chair of Management at the LINKS Center for Social Network Analysis, Gatton College of Business and Economics, University of Kentucky. He received his Ph.D. in 1989 from the University of California, Irvine. His research interests include social network theory and methodology, knowledge management, and social cognition.

**Notes**
The authors are grateful to the Connections editorial team and members of the LINKS Center for social network analysis at the University of Kentucky, especially Dan Brass, Joe Labianca, and Ajay Mehra for their help in shaping our thinking. This work was funded in part by a grant from DARPA (DARPA-SN-10-36: Graph Understanding and Analysis for Rapid Detection).

*Please address correspondence to Dan Halgin at danhalgin@uky.edu*

## 1. Introduction

In this article we introduce E-NET (Borgatti 2006), a software package specifically designed for the analysis of ego network data (as opposed to "whole-network" data), and in particular ego network data collected via a personal network research design (PNRD[1]). In presenting the capabilities of the program, we review key ideas in the analysis of ego network data, and discuss specific measures used to describe the size, composition, and structure of personal networks. We close with new directions in the analysis of such data focusing on longitudinal analysis.

An ego network consists of a focal node ("ego"), together with the nodes they are directly connected to (termed "alters") plus the ties, if any, among the alters, as shown in Figure 1. These networks

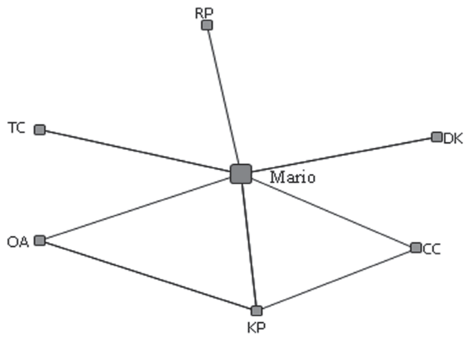



*Figure 1*. Ego network of Mario

are also known as personal networks, ego-centric networks, and first order neighborhoods of ego (cf. Everett & Borgatti, 2005; Mitchell, 1969; Wellman, 1979). Ego networks may be obtained by extracting them from a full network, as illustrated in Figure 2 (Holly's ego network is extracted from the full network). In that case (that is, when the full network is available), the decision to analyze just the ego networks is a theoretical choice to focus on the local rather than the global.

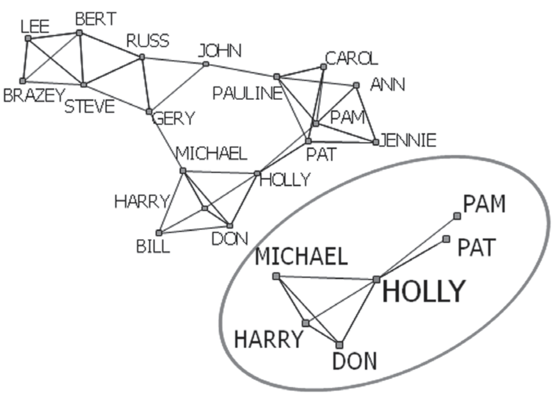



*Figure 2*. Extraction of Holly's ego network from a full network

Ego networks may also be collected directly, using a personal network research design or PNRD. A PNRD involves sampling a collection of unrelated respondents (called egos) and asking them about the people in their lives (called alters). For example, if we are interested in the social factors that influence entrepreneurial success, a personal network research design would involve sampling a set of unrelated entrepreneurs and ask each one about the resources that they derive from their personal contacts. We could easily interview entrepreneurs in different countries and relate aspects of their networks with some chosen dependent variable such as firm performance or funds collected. Although we sacrifice the ability to analyze global network measures, the personal network approach allows us to investigate whether successful entrepreneurs tend to have a greater number of contacts than others, whether entrepreneurs in New York tend to have a more diverse set of personal networks than those in Rome, or whether male entrepreneurs tend to have more personal contacts who run in different social circles than female entrepreneurs. We might also use the personal network approach to conduct an in-depth analysis of one focal entrepreneur.

In this example, it seems like the PNRD has advantages over the alternative design, the full network research design (FNRD). In a full network research design, we begin with a set of nodes, then measure all of the ties of a given type among those nodes. So in the study of entrepreneurs, the first design decision would be to determine which nodes to study. A natural first attempt would be to take as our nodes a set of entrepreneurs, such as all entrepreneurs belonging to a given association of entrepreneurs. But the people from whom entrepreneurs obtain key resources need not be other entrepreneurs. They could be friends, family, rich widows, and so on, who would probably not be in our population. A second attempt might be to study the whole of a small community in which there were entrepreneurs but also sources for these entrepreneurs, and we would measure ties among all pairs of nodes. This would yield very rich data, but could get enormous very fast, and much of the data might not be directly relevant to our research question. If we also wanted to compare New York and Rome entrepreneurs, we would have double the size of our study.

In summary, sometimes we will find it useful to collect our ego network data using a PNRD rather than an FNRD. If so, however, we will find that the data resulting from PNRDs are quite different from those generated by FNRDs, and computer programs designed to analyze full network data, such as UCINET (Borgatti Everett & Freeman 2002) will have significant difficulty coping with the FNRD data. This is the need filled by specialized PNRD software such as E-NET, presented here, and other programs such as EgoNet (McCarty 2003).

## 2. Data Collection in Personal Network Research Designs

In the personal network research design (PNRD), researchers collect network data by sampling unrelated and anonymous respondents from a large population and gathering information about each of their ego networks. A personal network survey can be administered by interviewers, or given to respondents to fill out, either on paper or via online methods.

The typical first step of the PNRD is to generate an exhaustive list of alters with whom the respondent has some type of relation-

[1]The personal network research design (PNRD) is defined later in the article.

ship. Termed a name generator, the respondent might be asked to list alters who occupy certain social roles (e.g., neighbors, kin, friends, coworkers), those with whom he shares interactions (e.g., discuss important matters with, has sex with, etc.), or those with whom he exchanges flows (e.g., borrowed money from, provide emotional support to). This approach is used in many classic studies of personal networks (e.g., Burt, 1984; Fischer, 1982; Laumann, 1966, 1973; Wellman, 1979). The open-ended nature of name generators can result in lengthy surveys so scholars should be aware of order-effects, fatigue, satisficing, non-redundancy, as well as interviewer effects (Marsden, 2003; Pustejovsky & Spillane, 2009; Van Tilberg, 1998). If faced with time constraints the researcher might limit the number of alters that each ego can nominate: "If you look back over the last six months, who are the four or five people with whom you discussed matters important to you?" (Burt, 1984)[2]. Or, the researcher might focus the name generator question to best match the specific line of research: "Among the people with whom you work, who has provided you with emotional support in the past six months?" Other name generator approaches include the reverse-small-world instrument (Killworth & Bernard, 1978), the phone book technique (e.g., Pool & Kochen, 1978; Freeman & Thompson, 1989), the modified multiple generator (MMG) and multiple generator random interpreter (MGRI) (Marin & Hampton, 2007), as well as techniques that incorporate visual interfaces (e.g, Kahn & Antonucci, 1980; McCarty & Govindaramanujam, 2006).

After obtaining a list of names using name generator questions, the researcher then typically asks the respondent name interpreter questions. These questions elicit additional information about ego's perceptions of the attributes of each alter (e.g., sex, race, income, etc.) and the shared relationship (e.g., duration, intensity, frequency, etc.). See Figure 3 for an example of a name interpreter grid. Name interpreter questions are unique to the PNRD and further highlight the ego-centered nature of this approach (as opposed to a full network design). Specifically, it is ego (not the alters) who provides information about the attributes of each alter. Researchers using a pure PNRD do not contact nominated alters to confirm alter attribute and relationship data and frequently use an alter-naming typology that allows ego to differentiate among alters without identifying them too closely (e.g., initials, code names, first three initials of first name and last name, etc.). This reduces privacy issues such as the lack of anonymity of alters in the full network approach and can be more favorable for respondents[3]. In addition, the key focus of personal network research is the ego-centered world of each respondent and how she views her alters (i.e., the number of alters, alter attributes, relationship attributes, the presence of relationships among alters, etc.). One limitation of the PNRD approach is that we do not check the accuracy of ego's view of his network (e.g., whether ego actually has ties with the nominated alters). In addition, we cannot fully determine the availability of all possible alters in ego's world. With full networks we have a bounded study population and can thus analyze whether certain individuals have more ties than expected to alters with certain attributes, say members of the same gender. With personal networks, we only see the world through ego's eyes so we don't know who ego chooses not to connect with. Consider the case where one of our entrepreneurs almost exclusively named alters who are male. The personal network approach does not allow us to determine whether the respondent consciously chooses to avoid women or if he lives in a community (e.g., Silicon Valley) with a highly unbalanced gender ratio and therefore has limited opportunities to nominate women.

Depending on the research goals, the researcher might also ask name interrelator questions that require the respondent to indicate whether the nominated alters themselves are connected. Due to time constraints and to avoid respondent fatigue, name interrelators typically use a reduced set of alters from the name generator (e.g., 10 alters) and one specific alter-alter relationship such as whether or not the alters know each other. Figure 4 is an example

| | **RP** | **CC** | **OA** | **TC** | **KP** |
|---|---|---|---|---|---|
| **RP** | | | | | |
| **CC** | 0 | | | | |
| **OA** | 0 | 0 | | | |
| **TC** | 0 | 0 | 0 | | |
| **KP** | 0 | 1 | 1 | 0 | |

*Figure 4*. The interrelator matrix for Mario's ego network

| Ego ID | Alter ID | Alter Age | Alter Gender | Alter Religion | Alter Income | Alter Frequency of Contact |
|---|---|---|---|---|---|---|
| Mario | RP | 32 | Male | Muslim | 55000 | 4 |
| Mario | CC | 18 | Female | Catholic | 23000 | 1 |
| Mario | OA | 28 | Female | Catholic | 64000 | 4 |
| Mario | TC | 56 | Male | Protestant | 43000 | 2 |
| Mario | KP | 31 | Male | Muslim | 17000 | 2 |

*Figure 3*. Name interpreter grid for ego Mario

[2] For a more detailed discussion of personal network size and how the number of alters can influence analytic options see the work of Killworth (et al., 1990), and McCarty (et al., 2001, 2002).

[3] The researcher might also consider a partial network research design using a snowball sampling method of contacting alters nominated by ego. See Marsden (2011) for an additional discussion of research design and network survey methods.

of a name interrelator matrix that might be given to the respondent. The alter identifiers in the matrix (in this case, initials) must be consistent with the identifiers of a set of alters that ego nominated in the name generator.In summary, there are pros and cons of the PNRD as presented in Table 1. For additional discussion of issues associated with the PNRD see Borgatti, Everett and Johnson (2013).

**Table 1**. Pros and Cons of PNRD

| PROS | CONS |
|---|---|
| • Able to sample respondents at random and generalize to well-specified populations. | • Difficult to determine the availability of possible alters in ego's world. |
| • Studies are scalable. Data storage and computing time requirements increase essentially linearly with the number of respondents, unlike whole network designs. | • Data contain only nominations by ego -- the choice to not nominate someone is not observed. |
| • Allows anonymity of respondents and alters, reducing privacy issues and promoting response rates. | • Alter-alter ties are as perceived by ego, and may be inaccurate. |
| • Data do not have statistical issues such as lack of independence that must be addressed when analyzing full network data. | • Inability to confirm ties (e.g., by looking at alter's response about ego). |
| | • Asking ego to report on every alter can be very time intensive. |

**3. Analysis of Personal Network Data Using E-NET**

Once collected, personal network data can be organized and analyzed using E-NET (Borgatti 2006), a free program that specializes in the analysis of ego network data, particularly data obtained via personal network research design. E-NET accepts data pertaining to egos (e.g., age, sex), alters (e.g., relationship between ego and alter, alter attributes), and relationships among the alters (e.g., whether ego 1 reports that alter A is connected with alter B). Other tools for the analysis of ego network data include VennMaker (Gamper et al., 2012) and EgoNet (McCarty 2003). VennMaker uses a visual interface that allows ego to move alters around the screen to elicit both the types of relations and their strengths. EgoNet uses a computer interface that elicits ego data from respondents using questions that achieve both name generation and interpretation. The program allows for visualization of ego networks, data management in a SPSS format, and a suite of standard network measures. In addition, Sciandra, Gioachin, and Finos (2012) provide an R package specific to the analysis of ego network data, and Müller, Wellman, and Marin (1999), provide guidelines for analyzing ego network in SPSS (For a discussion of additional options for the analysis and visualization of ego network data see McCarty et al., 2007).

*3.1 Importing Personal Network Data to E-NET*

Personal network data can be imported into E-NET in two formats, which we term row-wise, and column-wise. E-NET also reads full network data.

*3.2 Row-Wise Format*

In the row-wise format, data are recorded as three matrices corresponding to ego, ego-alter relationships, and alter-alter relationships[4]. The first matrix contains the collected attributes about each ego surveyed. The rows correspond to egos and the columns correspond to collected attributes about each ego. In the sample matrix provided in Figure 5, Ego 2 is a 30 year old woman with an income of 85000.

| ID | Age | Sex | Income |
|---|---|---|---|
| 1 | 21 | Male | 18000 |
| 2 | 30 | Female | 85000 |
| 3 | 45 | Female | 32000 |

*Figure 5*. Ego data matrix

The second matrix (see Figure 6) contains information about ego-alter relationships. These data include attributes of ego's relationship(s) with each alter (i.e., presence/absence, strength, frequency, etc.) as well as ego's perception of each alter. Note that each row of the matrix corresponds to ego's relationship with a unique alter. Egos with multiple alters will have multiple rows of data. If the researcher collects data about multiple kinds of relationships with each alter, the additional relationships are captured by added columns. In the sample matrix, Ego 1 has both a friendship and mentorship tie to alter 1_1, a friendship tie to alter 1_2, and a mentorship tie to alter 1_3. For convenience, we have used labels for the alters that include a reference to the ego they are attached to, hence alter "1_3" is the third alter of ego 1. The alter data matrix also includes the attributes of each alter (from ego's point of view). For example, alter 1_1 is a 40 year old woman, at least according to ego.

| From | To | Friend | Mentor | Alter Age | Alter Sex |
|---|---|---|---|---|---|
| 1 | 1_1 | 1 | 1 | 40 | Female |
| 1 | 1_2 | 1 | 0 | 33 | Male |
| 1 | 1_3 | 0 | 1 | 42 | Female |
| 2 | 2_1 | 0 | 1 | 63 | Male |
| 3 | 3_1 | 1 | 1 | 43 | Female |
| 3 | 3_2 | 1 | 1 | 21 | Female |

*Figure 6*. Alter data (row-based format)

[4] We recommend that users first create the three matrices in Excel (or a similar program).

As shown in Figure 7, the third matrix contains information about ego's perceptions of the presence of relationships (if any) among alters. The first two columns are the IDs of the alters that have a tie. For example, rows 1 and 2 of Figure 7 indicate that Ego 1 reports that Alter 1(1_1) knows both Alter 2 (1_2) and Alter 3 (1_3).

| From | To | Knows |
|---|---|---|
| 1_1 | 1_2 | 1 |
| 1_1 | 1_3 | 1 |

*Figure 7.* Alter-alter data (row-based format)

When ready for importation into ENET, the user compiles the three matrices into a single text file using the VNA format (See Figure 8 for an example). The VNA file can be created by copy-and-pasting the matrices from a spreadsheet editor (e.g., Excel) into any text editor (e.g., Notepad) and saving the document with file extension .vna. Note that in the VNA file the three kinds of data are identified by an asterisk and matrix title ("*ego data", "*alter data", "*alter-alter data").

```
*ego data
ID    Age   Sex      Income
1     21    Male     18000
2     30    Female   85000
3     45    Female   32000

*alter data
From  To    Friends  Mentor  Age   Sex
1     1_1   1        1       20    Female
1     1_2   1        0       33    Male
1     1_3   0        1       24    Female
2     2_1   0        1       63    Male
3     3_1   1        1       43    Female
3     3_2   1        1       21    Female

*alter-alter data
From  To    Knows
1_1   1_2   1
1_1   1_3   1
```

*Figure 8.* Row-based VNA file format

*3.3 Column-Wise Format*

E-NET also reads personal network data organized in an Excel file in what we term the column-wise format (see Figure 9). In this approach, the data are organized in one matrix such that each row corresponds to a specific respondent (ego) and columns correspond to ego attributes, ego-alter ties and perceptions, and alter-alter relationships. Note that the alter variables across the columns are repeated for each alter and labeled such that either the variable name is preceded by the alter number (e.g., A1Age, A2Age, A3Age, A1Sex, A2Sex, A3Sex) or vice versa (e.g., Age1 Age2 Age3 Sex1 Sex2 Sex3). Variables capturing ties among alters are named using the following format: "<variable name> <alter number> - <alter number>" (e.g., "knows1-2" indicates that alter1 knows alter 2). Using this naming convention enables E-NET to automatically identify ego variables, ego-alter ties and perceptions, and alter-alter ties. (If a different naming convention is used, the user can manually identify the variables by type within E-NET).

*3.4 Full Network*

E-NET also accepts full network data (i.e., sociometric data) stored in the UCINET data file format[5]. The analysis of full network data using E-NET is appropriate when the researcher is only interested in ego-centric measures. For instance, one might import the full network data presented in Figure 2 and use E-NET to extract, analyze, and compare the ego network of Holly with the ego networks of the other 17 individuals in the full network.

*3.5 Importing the Data*

Once the data are properly organized into a row-wise (vna file), column-wise (excel file), or full network data format (UCINET file), the appropriate file can be imported into E-NET under File | Open.

*3.6 Data Organization in E-NET*

In discussing the organization of the program, we use data from the 1985 General Social Survey (GSS) as a running example. The GSS is a personal-interview survey designed to monitor changes in both social characteristics and attitudes currently being conducted in the United States. In 1985, the survey included a selection of

| Age | Sex | Income | A1 Age | A2 Age | A3 Age | A1 Sex | A2 Sex | A3 Sex | A1 Friend | A2 Friend | A3 Friend | Knows 1-2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | Female | 18000 | 20 | 33 | 24 | Female | Male | Female | 1 | 1 | 1 | 1 |
| 30 | Male | 85000 | 63 | | | Male | | | 0 | | | 0 |
| 45 | Female | 32000 | 43 | 21 | | Female | Female | | 1 | 0 | | 1 |

*Figure 9.* Column-wise data format

[5] Please see the UCINET help guide for a detailed discussion of ways to import data into UCINET.

personal network questions based on the following name generator:

> From time to time, most people discuss important matters with other people. Looking back over the last six months-who are the people with whom you discussed matters important to you? Just tell me their first names or initials.

Name interpreter questions were used to elicit information about ego's relationship with each alter (e.g., intimacy, frequency of communication, type of relationship, duration) and ego's perceptions of the demographics of each alter (e.g., sex, religion, race, education, age, political views). Name interrelator questions were used to determine relationships among the alters. The dataset contains information on 1534 respondents and is organized in a column-wise format.

E-NET organizes its data into five key tabs visible in the main program window: Egos, Alters, Alter-Alter Ties, Visualization, and Measures.

*3.7 Egos Tab*

The Egos tab displays attributes of the egos. As we can see in Figure 10, The GSS dataset includes sex, race, education degree and income as respondent attributes. The middle right of the screen displays the number of records in the spreadsheet (1534), and the top right provides an area to enter SQL commands for filtering the data. For example, we could enter "sex = 'female' and race = 'white'" to restrict our analysis to respondents who were white women.

*3.8 Alters Tab*

The Alters tab (see Figure 11) displays the data obtained from the name interpreter, and consists of ego's perceptions of the alter attributes, together with the nature of the relationship between ego and alter. In the figure, we can see RCLOSE, which refers to how close the respondent feels to each alter, and AGE, SEX and EDUCATION, which refer to the age, sex and education of each alter, as reported by the respondent. In the grid, each row corresponds to a specific ego-alter. The first row in Figure 11 indicates that ego 0001 has a tie to alter 0001.1, and did not answer the question of how close the relationship was. But the respondent does indicate that alter 0001.1 is a 32-year old male college graduate. Rows 2 - 5 provide information about ego 0001's other alters.

*3.9 Alter-Alter Ties*

The Alter-Alter Ties tab (see Figure 12) displays data about relationships among alters. We can see in the first row of Figure 12 that ego 0001 has indicated that his alters 0001.2 and 0001.3 were "especially close". We can also see that the dataset contains 15,340 alter-alter ties.

*3.10 Visualization*

The Visualization tab (see Figure 13) lets the user visualize each ego network. The user can use the options to manually or automatically scroll through visualizations of the entered networks. When the user clicks on selected alters, the data chart on the left of the screen displays the associated relational data (e.g., attributes of the selected alter). In 0002's personal network visualization, information about alter 0002-2 (e.g., male, graduate of professional school education, white, Jewish) is displayed.

*3.11 Measures*

The Measures tab displays the output from the various analyses discussed in the next section. Figure 14 shows an example of output from the structural holes procedure. It should be noted that by clicking the Excel button on the toolbar, the user can transfer all the data in the measures tab to an Excel spreadsheet. The user can also transfer selected columns to the Egos tab, to be used as input for another analysis.

*3.12 Data Analysis using E-NET*

E-NET includes multiple analysis options appropriate for ego-network data. These are found under the ANALYZE menu option. The analysis choices fall into the general categories of composition and structure.

*3.13 Compositional*

We start with a discussion of compositional measures that focus purely on summarizing the characteristics of an ego's alters. These statistics are consistent with Lin's (1982) social resource theory, which is based on the resources an ego can access through relationships with different kinds of alters. For example, we might be interested in determining the distribution of alters in regards to a specific categorical variable such as sex, race, religion, level of education, etc. We might hypothesize that individuals who have ties to many women will have different views regarding gender roles than individuals who only have ties to men. Using the GSS dataset with E–NET, we can calculate the percentage and raw count of female alters in each respondent's network (found under Analyze|Composition). We can then use these statistics as variables in regression models that test our hypotheses.

We can also use E-NET to investigate the composition of alters in terms of continuous variables and calculate the average, maximum, minimum, total values, and standard deviations of selected alter attributes. For instance, we might hypothesize that egos with ties to alters who have a wide range of earning powers (measured by standard deviation of alter income) have different perspectives on financial issues than others, or that egos with elderly alters (measured by average age of alters) have different views on healthcare than those with younger alters. In the GSS data the average age of alters in collected ego networks range from 16.8 years (ego 1254) to 86 years (ego 0245).

We can also use E-NET to investigate the diversity of alters in each respondent's network with respect to specific variables (Under Analyze | Heterogeneity). For categorical variables, E-NET offers two classical measures of heterogeneity: Blau's index (also known as Herfindahl's measure and also Hirschman's measure), and Agresti's IQV (Agresti and Agresti, 1978). Egos whose alters

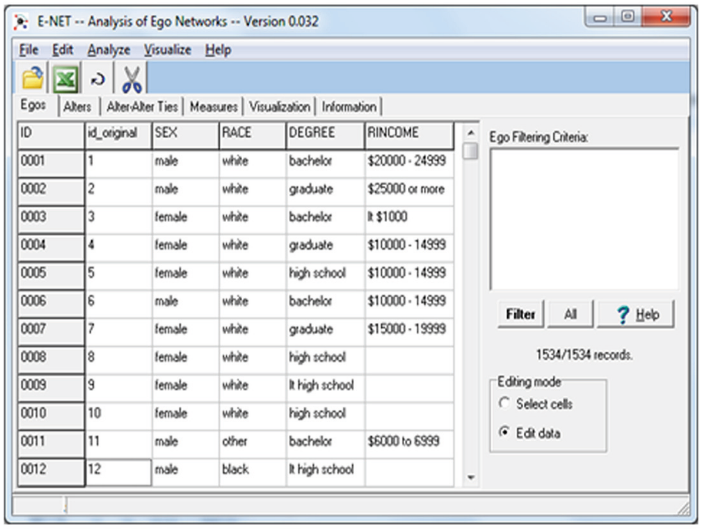



*Figure 10*. Egos tab in E-NET

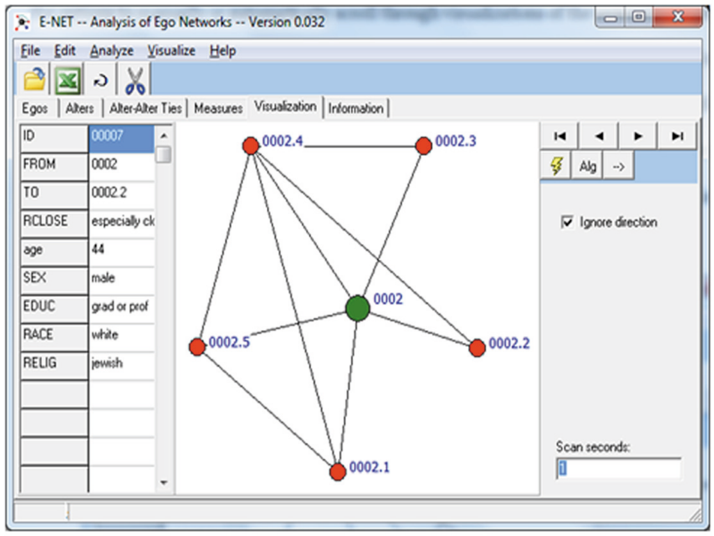



*Figure 13*. Visualization tab

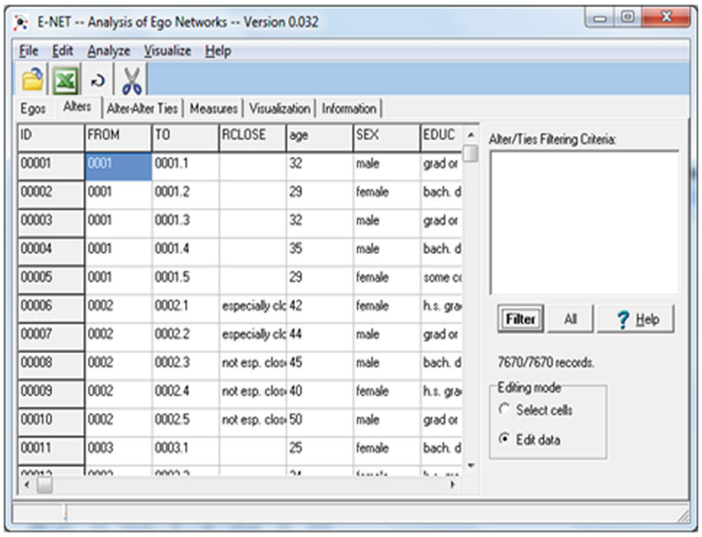



*Figure 11*. Alters tab

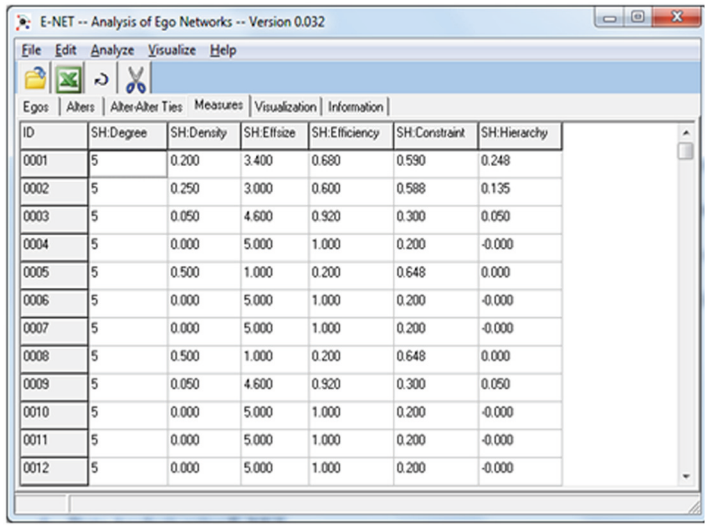



*Figure 14*. Measures tab

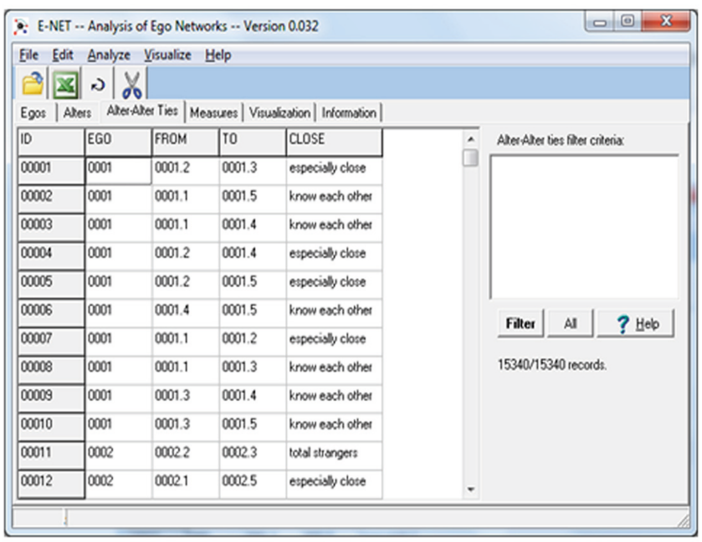



*Figure 12*. Alter-alter ties tab

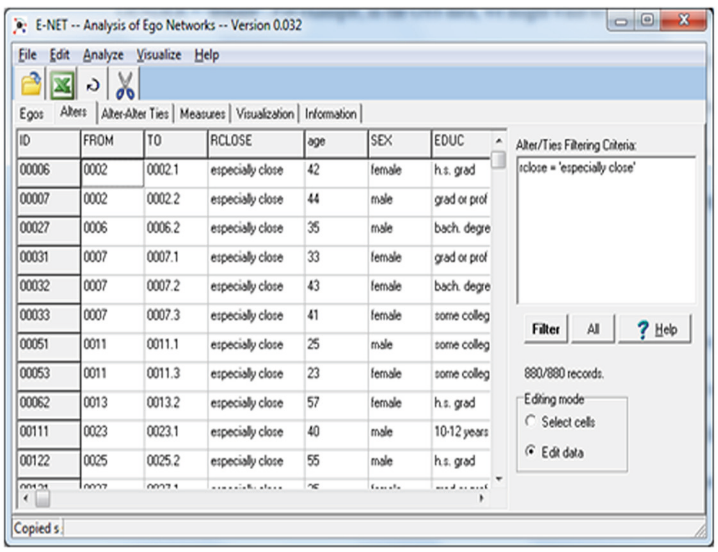



*Figure 15*. Filtering ties so that only "especially close" ties are allowed

are mostly the same with respect to some categorical attribute (e.g., gender or race), will have small heterogeneity scores while those with more diversity in their ego-networks will have a value closer to 1. For continuous variables such as age and income, E-NET computes the standard deviation of the alters' values.

E-NET also has the capability to construct aggregate cross-tabs of node attributes (found under the Analyze tab). Using the GSS dataset we can count the total number of personal network ties within and across gender categories (e.g., how frequently were men nominated by male egos?) or within and across racial categories (e.g., how frequently did Black respondents nominate White alters?). The crosstab command will also calculate chi square (and a p-statistic which is not adjusted for autocorrelation) and Yule's Q statistics. In addition, users are able to do crosstabs of alter attributes with other alter attributes. For example, we count the number of male alters that had ties with female alters.

Researchers can also use E-NET to investigate the similarity between ego and alters, termed homophily (e.g, Marsden, 1988; McPherson et al., 2001). Krackhardt and Stern's (1988) E-I statistic calculates ego's propensity to have ties with alters in the same group or class as self. The grouping variable is determined by the researcher. The measure is calculated by totaling ego's ties to alters who are "external" (i.e., those that are in a different attribute category), subtracting the number of ego's ties to alters who are "internal" (i.e., from the same attribute category) and dividing by network size. For example using the GSS data and race variable we note that ego 1007 self-identifies as Black and has four ties to alters who are Black and one tie to an alter of a different race category, resulting in an E-I score of - 0.6 (calculated as follows: (1 external – 4 internal) / (5 total) = - 0.6). Egos with ties to only those in the same selected category (e.g., ego is Hispanic and only has ties to other Hispanics) will have an E-I race score of -1 and those with only ties to those in other categories (e.g., ego is Hispanic and only has ties to alters who are White, Black, Asian, and categories other than Hispanic) will have an E-I race score of +1. This measure can be calculated for each individual ego, as well as the population as a whole. In the 1985 GSS data, the overall E-I score for race using "especially close" ties based on race is -0.895 indicating a strong preference on the part of egos for alters of the same race when it comes to strong ties.

### *3.14 Structural*

We can use E-NET to study structural characteristics of personal networks such as density or structural holes. According to structural holes theory (Burt, 1992), it is advantageous in many settings for ego to be connected to many alters who are themselves unconnected to the other alters in ego's network. Burt (2000) posits that individuals with networks with multiple structural holes are likely to receive more non-redundant information, which in turn can provide ego with the capability of performing better or being perceived as the source of new ideas. Networks rich in structural holes also provide ego with greater bargaining power and thus control over resources and outcomes, and greater visibility and career opportunities for ego throughout the social system (Burt, 1992, 1997; Seibert, Kraimer & Liden, 2001). These network structures have been shown to provide individuals with information that may prove useful to finding jobs (Granovetter, 1974), work place performance (Mehra, Kilduff & Brass, 2001), promotions (Brass, 1984, 1985; Burt, 1992) and creativity (Burt, 2004).

In the GSS dataset, some respondents have completely disconnected alters and thus maximum brokerage opportunities (represented by high effective size and efficiency, and low density and constraint), whereas others have alters who are all connected creating a closed personal network (indicated by low effective size and efficiency, and high density and constraint). Figure 14 shows the results of running structural holes on the GSS dataset.

### *3.15 Data Filtering*

As previously mentioned, a key feature of E-NET is its powerful filtering capability that allows the researcher to select which egos, which alters and which relations should be active in any given analysis. For instance, one might be interested in only analyzing the personal networks of white females, or only interested in considering ties to alters who are Hispanic, or only ties among alters that are especially close. Each of the data tabs (egos, alters, and alter-alter ties) has an area for entering filtering criteria. The criteria are entered using the SQL syntax used in database applications, such as AGE > 20 AND GENDER = 'female'. For example, in the GSS data, we might want to run analyses only on alters that ego indicates are "especially close." In that case, we would go to the alters tab, and type RCLOSE = 'especially close', as shown in Figure 15. We may also want to simply search for certain egos whose data we want to check. For example, we could enter RINCOME LIKE '%or%' to find respondents with incomes of "$250,000 or more" and "$18,000 or less".

## 4. Longitudinal Analysis

We close by describing new techniques for the analysis of ego-network data that use longitudinal data to analyze network change over time. The tools for doing this are currently being implemented in E-NET, and are already available (for full network data) in UCINET (Borgatti et al., 2002). To begin, let us consider an ego network analysis of the CAMPNET dataset available in UCINET. The original data were collected at two points in time at a research methods workshop in which 18 participants were asked to rank order the other participants based on the amount of time that they spent with each other in the previous week. For our purposes, it is convenient to dichotomize the data so that only the top three ranks are considered a tie. Although the data were collected via a full network design, we can use E-NET to extract and analyze each of the ego-networks that comprise the full network at both T1 and T2. Using these longitudinal data, we can analyze the pattern of changes in each actor's ego network.

Data of this type enable us to investigate questions such as whether men or women are more likely to make new friends, and whether low or high status actors are more likely to make relational choices that would be consistent with a strategy of increasing structural holes. A naïve approach to testing such hypotheses might be to calculate the appropriate personal network statistics (e.g., number of ties, number of structural holes) at each point in time, and compare the results to determine whether the genders were significantly different in the amount of increase in network size, or whether one status group added structural holes at a greater rate than the other. However, this approach fails in multiple areas. We highlight various shortcomings and introduce new methods that better address tie churn in personal networks.

Consider the simple case of making new friends. Suppose a given ego nominates three alters at T1 and three alters at T2, so that network size is 3 at both time periods. We can conclude that ego network size does not change, but we are not answering the original question. The obvious challenge is that we cannot determine whether, at T2, ego nominated the same three alters as at T1, or completely changed his network by dropping the initial three and finding three new alters.

Now consider the case of network diversity. Using the naïve approach, we measure Blau's heterogeneity index at T1 and T2. Suppose ego has ties to only women at T1. Blau's index will indicate a lack of heterogeneity in ego's network (heterogeneity = 0). If ego drops all T1 ties and forms all new relationships at T2 with alters who are all men, Blau's index will again indicate a lack of heterogeneity in ego's network (heterogeneity = 0). Clearly, this approach fails to capture major changes in composition. We might also look at changes in structure to identify egos who connect new pairs of alters over time. Again, we can calculate various structural measures and compare, but we won't be able to differentiate egos who make no changes in their networks from those who have dropped some ties and formed others, creating the same number of brokerage positions.

To better capture change in personal networks we propose more specific measures that separately measure the formation of new ties, the retention of existing ties, and the loss of old ties[6]-what collectively might be termed tie churn (Sasovova et al, 2010)[7]. Figure 16 displays selected tie churn statistics associated with ego networks in the CAMPNET dataset. Note that in this dataset, network size at T1 and T2 is, by design, consistent for all actors. However, some actors (such as Jennie, Michael, and Steve) keep the identical network over time as indicated by 0 new ties and 0 lost ties. Other actors have measures allowing us to test new hypotheses associated with change. For example, it can be argued from self-monitoring theory (Snyder, 1974) that high self-monitors are more able to add new ties, but because of a perception of inauthenticity, they will also be more likely to lose ties over time. We can test these two ideas with regressions relating self-monitoring score to number of ties added and number of ties dropped, respectively.

The tie churn approach can also be applied to the analysis of personal network composition. For instance, we might compare the attribute characteristics of the added, kept, and dropped alters to identify possible tie alteration strategies. Some egos might tend to form ties over time with alters of higher status, as exemplified in our CAMPNET data by Brazey, who dropped ties with fellow par-

| Actor | T1Size | T2Size | New Ties | Lost Ties | Kept Ties | New Ties to Instructors |
|---|---|---|---|---|---|---|
| HOLLY | 3 | 3 | 2 | 2 | 1 | 0 |
| BRAZEY | 3 | 3 | 2 | 2 | 1 | 2 |
| CAROL | 3 | 3 | 1 | 1 | 2 | 0 |
| PAM | 3 | 3 | 1 | 1 | 2 | 0 |
| PAT | 3 | 3 | 2 | 2 | 1 | 0 |
| JENNIE | 3 | 3 | 0 | 0 | 3 | 0 |
| PAULINE | 3 | 3 | 1 | 1 | 2 | 0 |
| ANN | 3 | 3 | 1 | 1 | 2 | 0 |
| MICHAEL | 3 | 3 | 0 | 0 | 3 | 0 |
| BILL | 3 | 3 | 1 | 1 | 2 | 0 |
| LEE | 3 | 3 | 1 | 1 | 2 | 0 |
| DON | 3 | 3 | 0 | 0 | 3 | 0 |
| JOHN | 3 | 3 | 1 | 1 | 2 | 1 |
| HARRY | 3 | 3 | 1 | 1 | 2 | 0 |
| GERY | 3 | 3 | 1 | 1 | 2 | 0 |
| STEVE | 3 | 3 | 0 | 0 | 3 | 0 |
| BERT | 3 | 3 | 1 | 1 | 2 | 0 |
| RUSS | 3 | 3 | 2 | 2 | 1 | 1 |

*Figure 16*. Selected tie churn statistics

[6] Feld and colleagues (2007) use a similar approach to investigate "which ties come and go."

[7] For full network data, we can also measure the retention of non-ties over time.

| | New Ties | Lost Ties | FoFa+ | FoFa- | T1 Holes | T2 Holes | Holes Added | Holes Lost |
|---|---|---|---|---|---|---|---|---|
| HOLLY | 2 | 2 | 1 | 1 | 2 | 3 | 5 | -3 |
| BRAZEY | 2 | 2 | 2 | 2 | 3 | 0 | 4 | -4 |
| CAROL | 1 | 1 | 1 | 3 | 2 | 1 | 0 | -2 |
| PAM | 1 | 1 | 1 | 2 | 2 | 1 | 1 | -1 |
| PAT | 2 | 2 | 0 | 0 | 1 | 3 | 5 | -1 |
| JENNIE | 0 | 0 | 0 | 0 | 1 | 2 | | |
| PAULINE | 1 | 1 | 1 | 1 | 2 | 1 | 0 | -2 |
| ANN | 1 | 1 | 1 | 1 | 2 | 1 | 2 | -2 |
| MICHAEL | 0 | 0 | 0 | 0 | 1 | 0 | | |
| BILL | 1 | 1 | 1 | 2 | 0 | 0 | 1 | 0 |
| LEE | 1 | 1 | 0 | 0 | 2 | 0 | 3 | -2 |
| DON | 0 | 0 | 0 | 0 | 1 | 0 | | |
| JOHN | 1 | 1 | 1 | 1 | 3 | 2 | 2 | -2 |
| HARRY | 1 | 1 | 1 | 2 | 1 | 0 | 1 | -1 |
| GERY | 1 | 1 | 0 | 0 | 0 | 2 | 3 | 0 |
| STEVE | 0 | 0 | 0 | 0 | 1 | 1 | | |
| BERT | 1 | 1 | 1 | 1 | 0 | 1 | 2 | 0 |
| RUSS | 2 | 2 | 2 | 2 | 2 | 1 | 4 | -2 |

**Note**: $^{+}$ New alters with ties to T1 alters, $^{-}$ New ties between new alters and T1 alters

*Figure 17*. Changes in ego network structure

ticipants in favor of ties with workshop instructors. Others might show a preference for connecting with alters who are similar to themselves. Virtually all of Holly's new ties were to women. Note that if ties dropped showed the same pattern, the pre- and post-compositions of the networks might look the same.

The tie churn approach also allows us to better investigate changes in structure in terms of triads and brokerage positions. Figure 17, displays different patterns of brokerage churn for these actors. We note that some actors experience churn that increases or decreases their brokerage opportunities. Pat replaced two ties and in the process added four new structural holes in her network; in contrast, Pauline added one new tie and dropped one tie with a net effect of losing two structural holes. Of course due to the full network nature of these data, actors also gain and lose holes as a result of the actions of their alters. For example, Don does not add or lose any ties (and therefore does not actively add or lose holes), but the actions of his alters result in the closure of his one hole at T1. Similarly, Bill makes a change that would potentially add one hole (as indicated by 1 in the Holes Added Column), but the hole is not realized due to changes of alters.

We can use this sort of micro-analysis to examine theoretical mechanisms of tie formation. For example, consider Russ in Figure 17. He replaced two alters between T1 and T2. The column labeled "Number of new alters with T1 ties to T1 alters" shows that both of the new alters had ties in T1 to the alter that remained. Thus, Russ tends to form new ties with friends of friends. In contrast, Pat formed two new ties with alters who did not have any existing ties to her T1 alters, indicating at the very least a willingness to form ties with people unconnected to her existing friends, perhaps reflecting a more entrepreneurial spirit.

At an even more detailed level, we can classify triadic structures that we see in ego networks, and examine rates of transition from one type to another. Given an Ego and potential Alter 1 and potential Alter 2, one method of classification yields six kinds of undirected triads: (1) Ego has ties to Alter 1 and Alter 2, and they have ties to each other (a closed triad, labeled E2A1 in Figure 18); (2) ego has ties to both alters, and they are not tied to each other (an open triad, often associated with brokerage opportunities, labeled E2A0); (3) Ego has a tie with Alter 1, who has a tie to Alter 2, but Ego has no tie to Alter 2 (E1A1); (4) Ego has a tie with Alter 1, and neither has a tie with Alter 2 (E1A0); (5) Ego has ties to neither potential alter, but they are tied to each other (E0A1); (6) none of the three nodes has a tie with another of the other three nodes (E0A0).

Applied to our CAMPNET dataset, Figure 18 provides an in-depth view of triadic change in Pat's personal network between T1 and T2. Note that at T1 Pat had two closed triads that no longer existed at T2 (in both cases Pat dropped ties with alters). In addition, Pat formed two new brokerage positions between T1 and T2 with the addition of ties to alters unconnected from each other. Even richer analyses are possible by taking into account attributes of the nodes, as suggested in a full network context by Gould and Fernandez (1989). For instance, we might determine whether Pat's changes place him as a broker between participants and workshop instructors, which likely has different benefits than being a broker between two participants.

Time 2

Time 1

| | E0A0 | E0A1 | E1A0 | E1A1 | E2A0 | E2A1 |
|---|---|---|---|---|---|---|
| **E0A0** | 54 | 5 | 17 | 2 | 1 | 0 |
| **E0A1** | 5 | 14 | 4 | 3 | 0 | 0 |
| **E1A0** | 20 | 1 | 12 | 2 | 2 | 0 |
| **E1A1** | 2 | 3 | 0 | 3 | 0 | 0 |
| **E2A0** | 0 | 0 | 1 | 0 | 0 | 0 |
| **E2A1** | 0 | 1 | 0 | 1 | 0 | 0 |

Legend:

**E0A0** = Null triad (no ties)
**E0A1** = Ego has no ties but the two potential alters are tied.
**E1A0** = Ego has tie to one alter; other potential alter is isolate.
**E1A1** = Ego has tie to one alter, who is tied to the other potential alter.
**E2A0** = Ego has ties to both alters, who are not tied to each other.
**E2A1** = Ego has ties to both alters, who are tied to each other.

*Figure 18*. Triad changes in Pat's personal network from T1 to T2

In summary, the tie churn approach to longitudinal personal network analysis, currently available in UCINET (Borgatti et al., 2002) and soon to be available in E-NET, reveals patterns that are not possible using a simple comparison approach. By considering new, kept, and dropped ties we can differentiate those who increase their brokerage from those who seek closure, those who seek diversity from those who seek homophily, and those who add ties to alters of higher status from those who add ties to alters of the same or lower status (e.g., Halgin et al., 2012; Sasovova et al., 2010).

**5. Conclusion**

Our principal goal in this article has been to present the basic steps of working with personal network research designs, from data collection through analysis, using a software package specifically designed for personal network research designs. We have also introduced some new measures for analyzing network churn. We hope that our discussion will help inform personal network analysis approaches and facilitate the generation of new network theory.

**References**

Agresti, A., & Agresti, B.F. (1978). Statistical analysis of qualitative variation. Pp. 204-207 in Sociological Methodology, eds. Karl F. Schuessler. San Francicso: Jossey-Bass.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Borgatti, S.P. (2006). E-NET Software for the Analysis of Ego-Network Data. Needham, MA: Analytic Technologies. https://sites.google.com/site/enetsoftware1/

Borgatti, S.P., Everett, M.G., & Johnson, J.C. (2013). Analyzing Social Network Data. Beverly Hills, CA: Sage.

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. (1985). Men's and women's networks: A study of interaction patterns and influence in an organization. *Academy of Management Journal, 28*, 327-343.

Burt, R.S. (1984). Network Items in the General Social Survey. *Social Networks, 6*, 293-339.

Burt, R.S. (1992). Structural Holes. The Social Structure of Competition, Cambridge, MA, Harvard University Press.

Burt, R.S. (1997). The contingent value of social capital. *Administrative Science Quarterly, 42*, 339-365.

Burt, R.S. (2000). The network structure of social capital. In Sutton, R., and Staw, B (Eds.) *Research in Organizational Behavior, 22*. Greenwich, CT: JAI press.

Burt, R.S. (2004). Structural holes and good ideas. American *Journal of Sociology, 110(2)*, 215 – 239.

Everett, M., & Borgatti, S.P. (2005). Ego network betweenness. *Social Networks, 27*, 31-38.

Feld, S.L., Suitor, J.J, & Hoegh, J.G. (2007). Describing changes in personal networks over time. *Field Methods, 19(2)*, 218-236.

Fischer, C.S. (1982). To Dwell Among Friends: Personal Networks in Town and City. Chicago : University of Chicago

Freeman, L.C., & Thompson, C.R. (1989). Estimating Acquaintanceship. In M. Kochen (ed.) The Small World. Ablex Publishing, New York: 147-158.

Gamper, M., Schönhuth, M., & Kronenwett, M. (2012). Bringing Qualitative and Quantitative Data Together: Collecting Network Data with the Help of the Software Tool VennMaker. In: Maytham Safar and Khaled Mahdi (eds.). Social Networking and Community Behavior Modeling: Qualitative and Quantitative Measures. Hershey (USA),193-213.

Gould, J., & Fernandez, J. (1989). Structures of mediation: A formal approach to brokerage in transaction networks. *Sociological Methodology, 19*, 89-126.

Granovetter, M. (1974). Getting a Job: A Study of Contacts and Careers. Cambridge: Harvard University Press.

Halgin, D.S., Gopalakrishnan, & G., Borgatti, S.P. (2012). A network perspective on social capital preferences: centrality, brokerage,and status aspirations. To be presented at the Academy of Management Conference, Boston, MA.

Kahn, D.A., & Antonucci, T.C. (1980). Convoys over the life course: Attachment, roles and social support, in Paul B. Baltes, and Orville G. Brim (eds) Life Span Development and Behavior, vol. 3. San Diego: Academic Press. 253-286.

Killworth, P. D., & Bernard, H.R. (1978). The reverse small-world experiment. *Social Networks, 1*, 159-192.

Killworth, P.D., Bernard,, H.R., McCarty, C., & Shelly, G.A. (1990). Estimating the size of personal networks. *Social Networks 12*, 289-312.

Krackhardt, D., & Stern, R.N. (1988). Informal networks and organizational crises: An experimental simulation. *Social Psychology Quarterly, 51(2)*, 123-140.

Laumann, E.O., (1966). Prestige and Association in an Urban Community. New York: Bobbs-Merrill.

Laumann, E.O., (1973). Bonds of Pluralism: The Form and Substance of Urban Social Networks. Wiley, New York.

Lin, N. (1982). Social resources and instrumental action in Social Structure and Network Analysis, edited by Marsden, P. and Lin, N. Beverly Hills, CA: Sage.

Marin, A., & Hampton, K.N. (2007). Simplifying the personal network name generator: Alternatives to traditional multiple and single name generators. *Field Methods, 19(2)*, 163-193.

Marsden, P.V., (1988). Homogeneity in confiding relations. Social *Networks, 10(1)*, 57-76.

Marsden,P.V. (2003). Interviewer effects in measuring network size using a single name generator. *Social Networks, 25(1)*, 1-16.

Marsden, P.V. (2011). Survey methods for network data. In John Scott and Peter J. Carrington (eds) The SAGE Handbook of Social Network Analysis. Sage, Los Angeles 370-388.

McCarty, C., (2002). Measuring structure in personal networks. *Journal of Social Structure, 3(1)*.

McCarty, C., (2003). EgoNet. Personal Network Software. Retrieved from http://sourceforge.net/projects/egonet/

McCarty, C., & Govindaramanujam, S. (2006). A modified elicitation of personal networks using dynamic visualization. *Connections, 26(2)*, 61-69.

McCarty, C., Killworth, P.D., Bernard, H.R., Johnsen, E.C., & Shelley, G.A. (2001). Comparing two methods for estimating network size. *Human Organization, 60(1)*, 28-39.

McCarty, C., Molina, J.L, Aguilar, C., & Rota, L. (2007). A comparison of social network mapping and personal network visualization. *Field Methods, 19*, 145-162.

McPherson, M., Smith-Lovin, L., & Cook, J.M., (2001). Birds of a feather: homophily in social networks. Annual Review of *Sociology, 27*, 415-444.

Mehra, A., Kilduff, M., & Brass, D. (2001). At the margins: a distinctiveness approach to the social identity and social networks of underrepresented groups. *Academy of Management Journal, 41(4)*, 441-452.

Mitchell, J.C. (1969). Social Networks in Urban Situations: Analyses of Personal Relationships in Central African Towns. Manchester, UK: Manchester University Press.

Müller, C., Wellman, B., & Marin, A. (1999). How to Use SPSS to Study Ego-Centered Networks. BMS: Bulletin of Sociological Methodology/Bulletin de Methodologie *Sociologique 64(1)*, 83-100.

Pool, I. deS., & Kochen, M. (1978). Contacts and influence. *Social Networks, 1*, 5-51.

Pustejovsky, J.E., & Spillane, J.P. (2009). Question-order effects in social network name generators. *Social Networks, 31*, 221-229.

Sasovova, Z., Mehra, A., Borgatti, S., & Schippers, M. (2010). Network churn: The effects of self-monitoring personality on brokerage dynamics. *Administrative Science Quarterly, 55(4)*, 639-670

Sciandra, A., Gioachin, F., & Finos, L. (2012). R package Egonet: A tool for egocentric measures in social network analysis. Available at http://cran.r-project.org/web/packages/egonet/egonet.pdf

Seibert S, Kraimer, M. & Liden, R. (2001). A social capital theory of career success. *Academy of Management Journal, 44*, 219-237.

Snyder, M. (1974). Self-monitoring of expressive behavior. *Journal of Personality and Social Psychology, 30*, 536-537.

Van Tilburg, T. (1998). Interviewer Effects in the Measurement of Personal Network Size: A Nonexperimental Study. *Sociological Methods and Research, 26*, 300-32.

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

# Welcome to the DEN! Data Exchange Network

As we strengthen Connections as an outlet for the scholarly work of the INSNA membership, we also want to maintain its special role as a vehicle for building and strengthening our community of network scholars. Toward this end, we are launching a new initiative, the Data Exchange Network (DEN), with a corresponding new section in Connections.

**DEN Mission**

The primary mission of DEN is to facilitate the sharing of instruments and network datasets among INSNA members. Obviously, many of us invest considerable time and effort in collecting data and may want to ensure that we have fully exploited the information before sharing it with others. However, we also understand that many of us have datasets that have either been poked and prodded beyond our desire to poke or prod any further, or that we find intriguing but not central to our research. In these cases, we may want to share them with others who may find them exciting in new and different ways. And, of course, many of us have data that we find so fascinating that, although we are still working on them, we are willing to share as a way of identifying researchers interested in the same phenomena, who may in turn become future collaborators.

Publishing in this section will generate benefits for the scholars who share their data, the INSNA community, and the field as a whole. In addition to building a shared resource that both experienced network researchers and newcomers to the field can draw upon and contribute to, maintaining a growing collection of vetted network data sets and instruments will help facilitate future research, education, and methodological advances. At the same time, members who donate their data or instruments will receive academic citations whenever they are used for publications. To accomplish this, submitted datasets and instruments must be accompanied by a brief article describing the instrument and/or the data and their collection. Publishing this article in the DEN section of Connections creates a citable reference to the original data or instrument.

Concurrent with publishing the article in Connections, the data and/or instrument will be uploaded and indexed on a section of the INSNA website, making them available to the whole membership. When other INSNA members use any of this data in their research or teaching, they will be expected to provide appropriate credit by citing the article in Connections that introduced the data. While the copyright to the original data will be retained by the scholar submitting it, uses of the data are governed by the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. In other words, data published in DEN can be freely used and modified for non-profit or scholarly research and teaching activities by INSNA members and their colleagues so long as any modified data are shared alike under the same license.

**Publishing in DEN**

To facilitate this process, we have developed several submission requirements and administrative procedures for the DEN section of Connections. Submissions must include a short article describing the data (not to exceed 2,500 words) and an electronic version of the network dataset and/or instrument being submitted.

The article, which describes the accompanying dataset and/or instrument, should follow standard conventions for submissions to Connections, specifying that the article is being submitted to the Data Exchange Network (DEN) section. (Guidelines for authors will be posted at http://www.insna.org/connections/DEN.html) The goal of the article is to help other INSNA members understand the data or instrument so that it may be effectively used in research or teaching. Even if only data are submitted, information about the questions used to collect the data should be provided. These articles need not be as detailed as a full codebook, but should provide enough detail that other members may appropriately use the data or measures. Naturally, if you want to supply the codebooks as part of the dataset to be published, we will be happy to include it. Additionally, the article should contain any information about the context from which the data were collected that may be relevant to others for appropriately using the data. If more than one relation, longitudinal data, or one or more attributes were collected in the same context, they should be bundled into one dataset and described in one article.

Data should be submitted in the most generic format possible (preferably in Excel) so that others are able to import them using a variety of software packages. Naturally, we are open to receiving data in other formats if Excel is impractical for some reason. Multiple relations, longitudinal data, and attributes may be presented on multiple worksheets within a workbook, or in multiple Excel files; however, as we want to share the data electronically amongst the membership, bundling as few files as possible will help facilitate downloading the data.

Consistent with the goal of providing citable references for these data and instruments, all materials submitted for the DEN will be peer-reviewed; however, unlike a more traditional review, the goal of the review process will be to ensure the utility and usability of the data/instrument by verifying that the data and methods are fully and clearly described and that any threats to validity are made transparent. Thus, the review process is designed to improve the documentation of the dataset and assess its potential utility to INSNA members, rather than to effect any direct theoretical or methodological contribution. Again, the principal goal is to build community and share resources through the Data Exchange Network, so the goal of the review process is to maximize the utility of the data and instruments by communicating as much relevant information as possible about the data, how they were collected, and any potential issues someone downloading the data should be aware of.

As an example of the type of article we hope to solicit for the DEN, we have provided an article describing the "Campnet" dataset distributed with UCINET as a template members can use for future submission. To maximize utility for others, descriptions of datasets submitted to this department should include the following sections:

I. Overview
A brief summary of the data. This is effectively an abstract for the dataset.

II. Data Collection
A narrative description of how data were collected, including a high-level description of questions asked or methods employed to collect the data.

III. Data Files and Formats
A description of the structure of the data accompanying this article. The data may be in any format useful to social network programs. In the accompanying template article examples of three such formats – full matrix, edgelist, and nodelist – are presented.

IV. Data Details
A required table detailing relevant information about the dataset, including:

a. response rates
b. non-respondent bias
c. any theoretical grounding for questions or methods employed
d. any existing publications employing these data
e. a short description of the *context*
f. nature of the *respondents*
g. whether the data are *longitudinal* and, if so, details about collection intervals
h. *temporality* of the data (e.g., the extent to which they are specific to the time at which they are collected)
i. *analytic utility* of the dataset – aspects others may find interesting in this dataset for teaching or research purposes.
j. known issues that threaten the validity of the data or anything else other INSNA members using these data for teaching or research should be aware of.

Italicized items will be used to provide a searchable interface in the electronic repository to help other members find datasets for particular teaching or research needs.

Submissions will be reviewed by at least two INSNA members.We are currently looking for reviewers to help with this process. If you are interested in reviewing for the DEN section of Connections, please contact either Rich DeJordy (r.dejordy@neu.edu) or Pacey Foster (pacey.foster@umb.edu), letting us know your interests and which tools you are most comfortable using for analysis. Finally, as the article describing the datasets will appear as a citable reference in the Connections journal, only datasets that meet appropriate institutional requirements for published research can be submitted for publication. In particular, authors must ensure all Ethics or Institutional Review Board requirements for human subject research have been met before submitting datasets involving human subjects.

In closing, it is our goal that the DEN become an easy way for INSNA members to share their data and instruments amongst themselves and contribute to the INSNA community by creating a valuable shared resource that provides appropriate academic credit to the contributors. We are very open to ideas from our fellow INSNA members about how to make this as beneficial as possible for the whole membership, so please feel free to contact the editor of Connections or send either of us a note if you have other suggestions, comments, or concerns about this new initiative.

Thanks!
*Rich DeJordy & Pacey Foster*
DEN Section Editors

# DEN
# Data Exchange Network

## The "Camp '92" Dataset

**Stephen P. Borgatti**
*University of Kentucky*

**H. Russell Bernard**
*University of Florida*

**Gery Ryan**
*Rand Corporation*

**Pertti Pelto**
*University of Conneticut*

**Rich DeJordy**
*Northeastern University*

### 1. Overview

The Camp '92 dataset is a simple, small dataset collected in 1992 at the National Science Foundation's Summer Institute for Ethnographic Research Methods in Anthropology. This was a three-week workshop involving 14 participants (13 faculty and 1 post-doc) and 4 instructors (3 faculty and 1 PhD student). The participants lived together at a motel for the duration. The motel had a common room where group members hung out. Subsets of group members regularly ate together.

The data consist of one kind of network tie, collected at two points in time, along with a limited set of individual characteristics. It includes three relations and three attributes. In two of the three relations, individuals rank ordered each other based on how much time they spent together (once in the second week of the workshop, and once again in the third week). In the third relation, only the top three people each participant spent time with are included in the relation. In addition, gender, role (Student or Instructor), and position (Faculty, Post-Doc, and PhD Student) were also recorded.

### 2. Data Collection

Respondents were presented a deck of 18 randomly shuffled cards representing the 18 members of the group. Respondents were asked to find the card with their name on it, and put it face down on a table. They then found the card representing the person with whom they spent the most time during the last week, and placed that card on top of the last one. This was repeated for all members of the group, ending with the card corresponding with the person they interacted with the least. This data collection procedure was repeated at the end of each week of the 3-week workshop (but the first week's data were lost, so only the 2nd and 3rd week's data are available).

The data were entered as a rank order data matrix such that, for each respondent, the person they said they spent the most time with was coded as a 1, the next person as a 2, incrementing through the person they said they spent the least time with coded as a 17. A 0 was used for the diagonal.

### 3. Data Files and Formats

All the data are provided in one Excel Workbook, called **Camp92 Dataset.xlsx**, containing 4 worksheets (tabs).

Three worksheets contain the relational data. For the purposes of introducing various possible data formats, we provide each network file in a different format, though typically relational data will use only one of these formats.

- Week_2. Valued rank-order data are provided in a full matrix format. The value of the cell (i,j) indicates the rank that respondent i assigned to person j. A value of 1 indicates that i reported spending more time with person j than with any other person. A value of 17 indicates that i reported spending less time with person j than with any other person.

- Week_3. Valued rank-order data are provided in edgelist format. The edgelist format includes one line of three columns for each pair of nodes. The first column indicates the respondent, the second column indicates the person the respondent responded about, and the third column indicates the value the person in the first column assigned to the person in the second column for the relation. In these data, a row with the values i, j, and k means respondent i said s/he spent the kth most time with person j.

- Campnet. This is a dichotomized version of the Week_3 matrix, such that all values less than or equal to 3 were recoded to 1, and all other values were recoded to 0. Thus, this matrix gives the top 3 interaction partners for each respondent. These data are presented in nodelist format. This format, which can only represent dichotomous data, is a multicolumn format where the first column is the respondent and all subsequent columns are people nominated by the respondent in response to a dichotomous relationship (e.g., "Is this person one of your top three interaction partners?"). Thus, the row i, a, b, c means respondent i said people a, b, and c were his/her top 3 inter action partners, in no particular order.

- Campattr. As these are attribute data, they are also presented in full matrix format, where cell (i,j) is respondent i's value for attribute j. The following three attributes are provided, forming an 18x3 matrix:

    - *Gender.* The gender of each person, where 1 = female and 2 = male.

    - *Role*. The person's part in the workshop, where 1 = participant and 2 = instructor.

    - *Position*. The person's academic position, where 1 = faculty, 2 = post-doc and 3 = phd student.

## 4. Data Details

| | |
|---|---|
| Response Rate | 100% |
| Non-Respondent Bias | N/A |
| Theoretical Grouping | These data were collected in a methods workshop specifically to illustrate the methods. So, no real theoretical objectives. |
| Publications Using These Data | • Borgatti, SP, Everett, MG, Johnson, JC. 2013. Analyzing Network Data. Sage Publications.<br>• Data are also distributed with UCINET software package as datasets camp92, campnet, and campattr. |
| Data Context | Research methods workshop/retreat |
| Respondents | Social science faculty and graduate students |
| Longitudinal | Yes, two points in time one week apart |
| Temporality | Low. Nothing about the data, collection, or context suggests the validity of the data will attenuate over time. |
| Analytical or Pedagogical Utility | • Demonstrating how valued graphs can be dichotomized<br>• Illustrating homophily with categorical attributes<br>• Finding cohesive subgroups<br>• Contrasting with core/periphery network structures<br>• Limitations of rank-ordering as a measurement method |
| Known Issues | None |

**Connections** publishes original empirical, theoretical, tutorial, and methodological articles that use social network analysis. The journal publishes significant work from any domain that is relevant to social network applications and methods. Commentaries or short papers in response to previous articles published in the journal are considered for publication. Review articles that critically synthesize a body of published research are also considered, but normally are included by invitation only. Book reviews, network images, review articles, and letters to the editor are also welcome. Manuscripts selected for publication are done so based on a peer-review process.

# INSNA

## International Network for Social Network Analysis

## International Network for Social Network Analysis

Connections is the official journal of the **International Network for Social Network Analysis** (INSNA). INSNA is a scientific organization made up of scholars across the world. Updated information about INSNA's annual conference (**Sunbelt Social Network Conferences**) can be found on the website at *www.insna.org*.



**INSNA (International Network for Social Network Analysis)**

Hardcopy circulation of Connections is sent to all members of INSNA, the International Network for Social Network Analysis, which has over 1300 members. Subscription to Connections can be obtained by registering for INSNA membership through the website: www.insna.org. Standard annual membership fee is US$60 ($40 for students). Wherever possible, items referenced in articles (such as data and software) are made available electronically through the INSNA website. In addition, the website provides access to a directory of members' email addresses, network datasets, software programs, and other items that lend themselves to electronic storage.

**Sunbelt Social Network Conferences**

Annual conferences for INSNA members take place in the United States for two years and in Europe every third year. The Sunbelt Conferences bring researchers together from all over the world to share current theoretical, empirical and methodological findings around social networks. Information on the annual Sunbelt Social Network Conferences can also be found on the INSNA website. Sunbelt XXXIII will be held in Hamburg, Germany May 21 - 26, 2013.

**Manuscript Submissions**

Submit articles to editorconnections@gmail.com. Manuscripts should be submitted as an MS Word document, and should not exceed 40 pages including all tables, figures, and references. All images, figures and tables should be sent as seperate files. Rastor (photographic) images and figures should be sent in a high resolution (300ppi min.) graphics format (EPS, TIFF, or JPEG), while line work images and figures should be sent as a vector-based format (EPS or SVG). Format and style of manuscript and references should conform to the conventions specified in the latest edition of the Publication Manual of the American Psychological Association. Further instruction on submission formatting can be found on the INSNA website, http://www.insna.org. Manuscripts that do not follow submissions criteria will be returned for revision. The journal follows a double-blind review process for research articles. 

**DEN (Data Exchange Network)**

In this issue, we are excited to introduce a new section in Connections, the Data Exchange Network section. The DEN feature has been designed with two goals: first is to build a community resource for network datasets and the second is to provide a format for citable references of datasets and instruments. Submissions should include an electronic version of the network dataset and/or instrument and a short article (not to exceed 2,500 words) describing the data being submitted. All materials submitted for the DEN will be peer-reviewed to ensure the utility and usability of the data/instrument. Accepted DEN contributions will appear in the hard copy of Connections, and the data sets will be available on the INSNA website through an indexed, searchable web interface. For more information, check out the introduction from the Section Editors in this issue.