Scientometrics (2013) 94:755–775
DOI 10.1007/s11192-012-0764-z

Examining the relationship of co-authorship network centrality and gender on academic research performance: the case of chemistry researchers in Pakistan

Kamal Badar · Julie M. Hite · Yuosre F. Badir

Received: 21 February 2012 / Published online: 12 May 2012

Abstract This research examines the association of co-authorship network centrality (degree, closeness and betweeness) and the academic research performance of chemistry researchers in Pakistan. Higher centrality in the co-authorship network is hypothesized to be positively related to performance, in terms of academic publication, with gender having a positive moderating effect for female researchers. Using social network analysis, this study examines the bibliometric data (2002–2009) from ISI Web of Science for the co-authorship network of 2,027 Pakistani authors publishing in the field of Chemistry. A non-temporal analysis using node-level regression reports positive impact of degree and closeness and negative impact of betweeness centrality on research performance. Temporal analysis using node-level regression (time 1: 2002–2005; time 2: 2006–2009) confirms the direction of causality and demonstrates the positive association of degree and closeness centrality on research performance. Findings indicate a moderating role of gender on the relationship of both degree and closeness centrality with research performance for Pakistani female authors.

Keywords Co-authorship network · Research performance · Network centrality · Gender · Social network analysis · Non-temporal and temporal analysis · Node-level regression

K. Badar
Institute of Management Sciences (IMS), University of Balochistan, Sariab Road, Quetta, Pakistan

K. Badar (✉) · Y. F. Badir
School of Management, Asian Institute of Technology, P.O. Box 4, Klong Luang, Pathumthani 12120, Thailand
e-mail: Kamal.badar@ait.asia; kamal.badar1980@gmail.com

J. M. Hite
Department of Educational Leadership & Foundations, Brigham Young University, 306 MCKB, Provo, UT 84601, USA

Introduction

It has long been realized that the co-authorship of articles in academic journals provides a window on patterns of collaboration within the academic community (Newman 2004). Co-authorship of a paper can be defined as documenting collaboration between two or more authors, and these collaborations form a *co-authorship network* which can comfortably be placed in the domain of social networks (Zurián et al. 2007; Gossart and Özman 2009; Newman 2001a, b, 2004; Nagpaul 2002) with authors being the nodes in the network and relationship between them established when they co-author a paper (Newman 2004). Such a network, might help raising both the quantity and quality of an individual's academic publishing record such that the researchers who are part of the network might benefit from the social capital which accrues to them because of their relationship with other researchers (Eaton et al. 1999; Jansen et al. 2010; Liao 2011; Mcfadyen and Cannella 2004; McFadyen et al. 2009; Oh et al. 2005; Yan and Ding 2009).

Social capital is the sum of actual and potential resources embedded within, available through and derived from the network of relationships possessed by an individual or social unit (Nahapiet and Ghoshal 1998) and thus comprises of both network and assets which can be mobilized through that network (Burt 1992). There are critical dimensions of social capital namely structural, relational and cognitive dimensions (Nahapiet and Ghoshal 1998). According to Nahapiet and Ghoshal (1998), all three dimensions are interrelated in important and complex ways and structural dimension of social capital often acts as a pre-requisite for relational and cognitive dimensions. In addition, the structural dimension of social capital highly overlaps with the concept of a social network (Zheng 2008).

This study examines centrality, an important facet of structural dimension of social capital (Bhardwaj et al. 2008; Zheng 2008) in co-authorship networks. Centrality specifies structurally advantageous positions of actors in the network (Bhardwaj et al. 2008). Previous research has studied the concept of centrality and its impact on innovation and performance in organizational networks (Tsai 2001; Sparrowe et al. 2001; Yang 2007), yet this concept has not been sufficiently studied in academic context. While few studies, have explored the concept of centrality in academic co-authorship networks (Zurián et al. 2007; Gossart and Özman 2009; Newman 2001a, b, 2004; Nagpaul 2002), they lack an insight about the outcomes of centrality for individuals in the network.

When studies have filled this gap by investigating and identifying the positive influence of centrality on performance outcomes in co-authorship networks (Abbasi et al. 2011; Eaton et al. 1999; Liao 2011), they leave some important research questions largely unanswered. First, *are there any differential performance benefits between different dimensions of centrality?* Second, "*are there any potential moderating effects of individual characteristics on the relationship?* These questions stem from emerging empirical evidence concerning: the conceptual distinction between various centrality dimensions (Valente et al. 2008), and role of individual characteristics in shaping payoffs from structural social capital (Bhardwaj et al. 2008; Burt 1998), respectively.

In addition the co-authorship literature does not demonstrate many studies of co-authorship networks in underdeveloped/developing countries. Yet these networks may provide potentially interesting cases due to the underdeveloped nature of research infrastructure and the lack of formal mechanisms for knowledge exchange and diffusion (Gossart and Özman 2009; Wagner 2006). While some research has studied co-authorship networks in Iran (Yousefi-Nooraie et al. 2008), Turkey (Gossart and Özman 2009) and India (Nagpaul 2002), all of these studies are conducted at the organizational level of analysis ignoring individuals, their centrality and their characteristics which might play an important role within the network. Thus the impact of

co-authoring network structure on performance may benefit from more evidence from developing/underdeveloped countries, particularly at the individual level of analysis.

Hence, this study, unlike others, attempts to extend the knowledge of co-authorship networks by focusing on a domestic co-authorship network of researchers publishing in Chemistry and it's sub-fields from a developing country (Pakistan), and examining the differential influence of three classical dimensions of centrality (degree, closeness and betweeness) on the research performance of each co-author. This study assumes that authors occupying positions high on each dimension of centrality (degree, closeness and betweeness) benefit differently from these central positions; hence enhancement of their research performance might be a possible outcome of their network centrality. Furthermore, unlike other studies, this research examines the moderating effect of an individual characteristic, gender, on the relationship between centrality and research performance. Thus these are the two important purposes of the study.

The remainder of the paper is organized as follows: Theory and hypothesis presents hypotheses development based on review of relevant literature, Method and Measures presents the methodology encompassing description about bibliometric network data and measurement of relevant variables, Data analysis techniques describe the techniques used to analyze data, Results and discussion presents the empirical results and their discussion and Conclusions and recommendations concludes.

Theory and hypothesis

Research performance

Research performance of academic researchers can be defined in terms of the quantity, quality and/or impact of their literary output. Objective indicators to assess research performance of academic researchers is frequently required by science managers, education policy makers and university administrators so that decisions can be made to regarding promotions, recruitments, award of grants and/or funds (Pike 2010).

A number of quantitative metrics have been proposed that (in principle) allow the comparison of individuals' scientific quality or impact (Cartwright and McGhee 2005; Cheek et al. 2006; Meho 2007), generally falling into the categories of reputation, yield or productivity, and influence or impact (Avital and Collopy 2001). Commonly used metrics include the total number of papers published, which is used to gauge basal productivity (Pike 2010); the mean or total number of citations received, which indicate the scientific utility of a study and can thus be used as a partial indicator of a study's quality and impact (Oh et al. 2005); and the journals where the papers were published and these journals' impact parameter (Liao 2011). In line with other studies, this study weights each publication by the journal's ISI impact factor,[1] to indicate research performance of each researcher (Liao 2011; Mcfadyen and Cannella 2004; McFadyen et al. 2009). Weighting a publication by impact factor allows evaluating the relative importance of the knowledge provided to the scientific community. The measure also eliminates any advantage of larger journals over smaller, frequently issued journals over less frequently issued and older journals over newer (McFadyen et al. 2009). In addition this particular indicator is particularly chosen in the country context of the study as it is officially used by the Higher Education Commission of Pakistan (HEC) to assess research

[1] The journal Impact Factor is the average number of times articles from the journal published in the past 2 years have been cited in the Journal citation report (JCR) year.

performance of researchers. For example, to be eligible to apply for the "approved PhD supervisor" a researcher must have an aggregate impact factor of five (Higher Education Commission of Pakistan, http://beta.hec.gov.pk/Pages/HECMain).

Centrality

Research on social networks has focused on the structural characteristics of social networks in explaining outcomes (Granovetter 1985); the structural dimension of social capital or simply structural social capital (Bhardwaj et al. 2008; Zheng 2008). The fundamental tenet of this perspective is that the location of an actor in the network enhances or constrains access to certain resources (Brass 1984; Ibarra 1993). Coleman (1990) noted that "having positions rather than persons as elements of the structure has provided one form of social capital that can maintain stability in the face of instability of individuals" (p. 320). In a structurally embedded relationship, the actor's location in a network of relations and interactions therefore provide benefits to that actor (Tsai and Ghoshal 1998) that may result in instrumental outcomes, for instance a firm's innovation (Ahuja 2000) and performance (Tsai 2001; Yang 2007) and an individual's creativity (Perry-Smith and Shalley 2003) and knowledge creation (Mcfadyen and Cannella 2004).

In explaining an actor's structural position in a network, centrality is most often applied as a property of social networks (Marsden 2002). One's centrality is defined as the extent to which a focal actor is connected to other actors within a specified network (Wasserman and Faust 1994). Centrality analysis, in fact is not new to sociology (Yan and Ding 2009). In a groundbreaking piece, Freeman (1977) developed a set of measures for centrality based on betweenness. In a follow-up article, Freeman (1979) proposed more dimensions of centrality in a social network, which have been developed into degree centrality, closeness centrality and betweenness centrality. This current study uses these three classical dimensions of centrality. Although considerable conceptual overlap exists between these dimensions, and they are often correlated, they are also conceptually distinct (Valente et al. 2008). For example, a node in the center of a star or wheel is the most central node in the network by all centrality measures (Freeman 1979). In other network configurations, however, nodes with high degree centrality are not necessarily the most strategically located. One way to characterize such distinctions among these concepts is in terms of how actors who occupy positions high on each dimension of centrality (i.e. degree, closeness and betweeness) might derive differential benefits from these positions (degree centrality providing benefits of direct ties, closeness providing benefits of quick flow of information and betweeness providing benefits of brokerage and control) and hence derive instrumental outcomes (hypotheses section provides more detail).

The three classical centrality dimensions can be mathematically defined as follows:

Degree centrality

Degree centrality equals the number of direct ties that an actor has with other actors or equivalently the number of actors directly connected to it. Thus degree centrality of a node n_i is mathematically given (Scott 1991):

$$C_{\mathrm{D}}(n_i) = \sum_{j=1}^{g} a(n_i, n_j) \qquad \left.\begin{array}{l} a(n_i, n_j) = 1 \text{ if and only if } n_i \text{ and} n_j \text{ are connected} \\ (n_i, n_j) = 0 \text{ if } n_i \text{ and } n_j \text{ are not connected} \end{array}\right\} \quad (1)$$

where g is the total number of nodes in the network and a (n_i, n_j) is a function which is equal to 1 if and only if node n_i and n_j are connected and zero otherwise. In order to compare networks of different sizes, normalized version of degree centrality has been proposed (Freeman 1979) which can be defined as the proportion of nodes adjacent to n_i:

$$C'_{\mathrm{D}}(n_i) = \frac{\sum_{j=1}^{g} a(n_i, n_j)}{g-1} \tag{2}$$

The normalized index can range from 0 to 1 (0 to 100 if expressed in percentage). It equals 1(100 %) when a node is directly connected to all others and 0 if it is isolated (not directly connected).

Closeness centrality

Closeness centrality focuses on how "close" an actor is to all other actors. It is measured as a function of mean geodesic/shortest distances. Suppose $d(n_i, n_j)$ is the length of geodesic path from n_i to n_j, meaning the number of edges along the path.[2] Then the mean geodesic distance from n_i to n_j averaged over all vertices g in the network (Newman 2010) is:

$$\ell(n_i) = \frac{1}{g-1} \sum_{j=1}^{g} d(n_i, n_j) \tag{3}$$

This quantity takes low values for vertices that are separated from others by only a short geodesic distance on average. Due to this reason, the mean distance $\ell(n_i)$ is not a centrality measure since it gives low values for more central vertices. Hence closeness centrality is calculated as an inverse of $\ell(n_i)$ and can be defined as inverse average distance between node n_i and all other nodes (Newman 2010).

$$C_{\mathrm{C}}(n_i) = \frac{1}{\ell(n_i)} = \frac{g-1}{\sum_{j=1}^{g} \mathrm{d}(n_i, n_j)} \tag{4}$$

Equation 4 above can range from 0 to 1 (0 to 100 if expressed in percentage). It equals 1(100 %) when a node is adjacent to all other nodes and 0 if one or more nodes is not reachable from node in question. Hence this index is only meaningful for connected network.

Betweeness centrality

Betweenness centrality measures the extent to which an actor lies on the geodesic paths between other actors. The important idea here is that an actor is central if it lies between other actors on their geodesics, implying that to have a large "betweeness" centrality, the actor must be "between" many of the actors via their geodesic. Suppose that a node n_j and n_k are connected in a network via several geodesic paths. Hence each geodesic is equally likely to be used. Let g_{jk} be the number of geodesics linking the two nodes. If a distinct node n_i lies on any of the geodesics linking n_j and n_k, we can label $g_{jk}(n_i)$ as the number of geodesics linking the two nodes that contain node n_i. $g_{jk}(n_i)/g_{jk}$ hence is the probability of node n_i lying "between" nodes n_j and n_k.

Therefore betweeness centrality for node n_i, is simply the sum of these estimated probabilities over all pairs of actors excluding n_i.(Wasserman and Faust 1994).

[2] Geodesic paths need not be unique i.e. nodes can be joined by several shortest path of same length. The length $d(n_i, n_j)$ however is always well defined, being the length of any one of these paths.

$$C_B(n_i) = \sum_{j<k} \breve{g}_{jk}(n_i)/\breve{g}_{jk} \quad i \neq j \neq k \tag{5}$$

Just like other measures, this measure also depends on g so it is also standardized like other centrality measures. Because we consider pairs of actors in this measure, we standardize it using $(g-1)(g-2)/2$ which is the maximum number of pairs of actors in a undirected network excluding n_i.(Wasserman and Faust 1994) Hence:

$$C'_B(n_i) = \frac{\sum_{j<k} \breve{g}_{jk}(n_i)/\breve{g}_{jk}}{(g-1)(g-2)/2} \tag{6}$$

Equation 6 above can range from 0 to 1 (0 to 100 if expressed in percentage). It is 1(100 %) when a node lies on all geodesics of all pairs of nodes and 0 when it lies on no geodesics.

Individual characteristics

This study argues that the relationship between centrality and research performance of individual researchers is not a straightforward proposition. Researchers with certain individual characteristics are expected to benefit more or less from the central position they have in the network. Limited evidence exists about the role of actor attributes on how they may benefit from structurally central positions. For example, Burt (1998) found that structural social capital and job promotion linkages were weaker for women managers, and Bhardwaj et al. (2008) found the relationship between structural social capital (network centrality) and satisfaction positive for whites but not for other racial minorities. This study proposes to examine the role of gender, a potentially moderating variable on the relationship between centrality and performance considering this as one of the purposes of the study. Gender is chosen specifically in Pakistan's context because Pakistan is a country where pervasive patriarchal attitudes and deep-rooted traditional and cultural stereotypes regarding the roles and responsibilities of women and men in the family, in the workplace and in society still prevail (See UN report: The World's Women, 2010).

Hypotheses

Degree centrality and research performance

Degree centrality equals the number of ties that an actor in a social network has with other actors (Freeman 1979). In other words it is the number of direct ties of an actor in the social network (Ahuja 2000). Direct ties can provide benefits namely knowledge sharing (Ahuja 2000; Berg et al. 1982) and joining of complementary skills (Ahuja 2000; Arora 1990; Richardson 1972). For instance if two or more authors co-author a paper, each contributes a certain amount of knowledge to the paper, therefore each author gains new knowledge through direct interaction and discussion among themselves(Eaton et al. 1999). If authors have same knowledge background, they benefit from bringing their own point of views to the topic, which helps in deepening the discussion (Abbasi et al. 2011). If authors have complementary knowledge, they benefit from learning each others' research and domain of expertise (Avkiran 1997). If authors have totally different knowledge background, they can enjoy each other's economies of specialization without investing in developing the specialization themselves (Ahuja 2000) and most likely produce new knowledge that resulted

from the combination of the two different knowledge backgrounds. This knowledge sharing and creation subsequently may result in higher research performance of each author (Abbasi et al. 2011; Eaton et al. 1999; Liao 2011). Therefore direct ties are expected to stimulate combination and exchange of resources within the relationships (Nahapiet and Ghoshal 1998) and provide researchers with access not only to new knowledge but also to new experiences and hence an increase in research performance might be a possible outcome.

Therefore we can propose the following hypothesis:

H1a Authors having higher degree centrality scores in the co-authorship network will have higher research performance (in terms of aggregate impact factor).

Closeness centrality and research performance

Degree centrality is a local measure and does not consider the indirect ties an actor has. Indirect ties tend to have their own benefits for network actors. For instance from a firm's perspective, a firm's partners bring knowledge and experience from their interaction with their other partners to their interaction with the focal firm (Gulati and Martin 1999). Therefore a firm can get access to not only the knowledge held by its partners but also held by its partner's partners (Gulati and Martin 1999). Very similarly, co-authors of an author can bring knowledge and experience from their own respective co-authors. Therefore a central author can have access to knowledge and skills of not only his/her immediate co-authors but to that of the co-authors of his/her co-authors as well.

Closeness centrality is a measure which considers ties of an actor with all other actors and is a so called global measure (Freeman 1979). As the name suggests, it specifies closeness i.e. how close an author is to rest of the authors. A high closeness centrality for an author specifies that they have the ability to access a larger portion of individuals in the network leading to a higher likelihood of exposure to various disparate authors and closeness to more clusters of highly connected authors (Perry-Smith and Shalley 2003). Hence, knowledge and complementary skills are likely to travel quickly and early (Borgatti 1995) to that author. This increased access for central authors means that chances of being exposed to knowledge and skills in the co-authorship network increases which may lead to enhanced research performance (Perry-Smith and Shalley 2003). Note that closeness centrality takes into account indirect ties but does NOT measure the number of indirect ties.

Therefore we can hypothesize that:

H1b Authors having higher closeness centrality scores in the co-authorship network will have higher research performance (in terms of aggregate impact factor).

Betweeness centrality and research performance

Apart from direct and indirect ties, actors who have ties connecting disconnected others or in other words, having ties spanning social divides tend to benefit from knowledge brokerage and control (Burt 2005). These actors tend to be gatekeepers of knowledge. In addition these actors tend to tap in non-redundant sources of knowledge and skills because disconnected actors might have diverse knowledge which the broker or gatekeeper can use to derive instrumental outcomes (Mehra et al. 2001).

Betweenness centrality is based on the number of shortest paths passing through an actor (Freeman 1979) or in other words extent to which author lies "in-between" other actors. In co-authorship networks, betweenness reflects how close the sub-network to which the author belongs is and how important the author's role as a broker is (Yan and Ding 2009). Thus,

betweenness creates advantage by lowering the risk of collaboration and by increasing the value of collaboration (Burt 2005). For instance an author co-authoring with two authors who themselves never co-authored, can benefit from diverse knowledge and skills of both (non-redundant knowledge). In addition, the disconnected co-authors depend upon the main author in order to exchange knowledge and skills (gate-keeping, brokerage or control of knowledge).

Hence betweenness centrality or the extent to which an author is on the shortest information path connecting other authors who are not themselves connected, provide brokerage opportunities to authors who can gain access to non-redundant and diverse pools of knowledge and skills and apply them to for their own sake to enhance their research performance.

Thus we can hypothesize that:

H1c Authors having higher betweeness centrality scores in the co-authorship network will have higher research performance (in terms of aggregate impact factor).

Gender and centrality

Past research suggests that gender has a strong impact on the development of social networks (Burt 1998; Brass 1985; Ibarra 1993, 1997) and women can be less central (Tharenou 1999) or more central (Brass 1985) in the networks as compared to their male counterparts. One exceptional study specifically studying co-authorship networks reports men and women having equally high central positions as female scientists and male scientists (Kretschmer and Aguillo 2005).

Gender is particularly important in a developing country like Pakistan which is predominately a male oriented society where issues of traditional negative gender-role stereotyping for women, sexual discrimination, domestic violence, fundamentalism and intimidation against women and lower literacy and employment rates for women are still prevalent (see UN report: The World's Women 2010). Hence it is expected that women in the labor force have to work in a traditional and sexually tense environment which might impose serious constraints on their performance. An immediate question that comes in mind here is whether *these constraints impact working/studying women in academia/research, specifically those who co-author in scientific publications and hence are part of the co-authorship network*. It can be implied that women who have advanced so far and risen to the levels of producing scientific output might have *defied some if not all odds*. They might have broken through some of the barriers. They might be the stronger, confident and exceptional sub-sample of a deprived population of females. The following arguments, stemming from the traditional and discriminatory nature of Pakistani society, specify how women might use co-authorship network ties, specifically, to break some barriers and hence derive stronger benefits of having a structurally central position in the co-authorship network as compared to men:

The presumed male domination of Pakistani organizations and numerical imbalance between genders (Fairhurst and Snavely 1983; Kanter 1977) most likely leads to great discomfort in cross-gender interactions (Hendrick 1981). More over there are some barriers unique to cross-gender interaction in Pakistani society, including the transfer of sex roles that are traditionally performed in a family context to workplace interactions, potential sexual tension, and public perceptions of a sexual relationship and sexual harassment (Kanter 1977; Kram 1988). In addition men as a typically dominant and top level (hierarchy) group in the Pakistani organizations and society might intentionally exclude women from the *club*, or *old boys'* network (Albrecht 1983) considering them as "illegitimate" members or outsiders (Burt 1998). Women employed/enrolled in academic/research institutions might overcome the barriers associated with cross-gender interaction to some

extent, by establishing co-authorship network ties. Women faculty at Pakistani Universities and research institutes might have male students as co-authors and/or male faculty might have female students as co-authors. In addition, male and female faculty, as colleagues might agree to co-author in scientific publications. Hence they might reap some benefits of cross-gender interactions such as establishing relationship with the dominant (men) and top level (hierarchy) group and thus *borrowing* some social capital (Burt 1998).

Working women in Pakistan face difficulties related to management of work and family (Ibarra 1993). Their responsibilities to shoulder the predominant share of domestic and child care activities (Ragins and Sundstrom 1989), may also imply a limit on the time that women have available (Kay and Hagan 1999; Seron and Ferris 1995) for employment related activities (see UN report: The World's Women 2010 for facts and figures). In addition these additional responsibilities might restrict women's mobility. Female academics/researchers on the other hand might break this barrier to some extent by establishing co-authorship network ties. Establishing these ties might lead to *division of labor* and hence women can still perform despite limited time available to them.

All of the above arguments lead to thinking that female researchers might be smarter in planning their co-authorship ties as compared to male researchers. In other words, female researchers might be more sensitive to and more aware of the structure of their co-authorship network (Brass1985). This might help shaping the structure of the co-authorship network in favor of female researchers. Male researchers, on the other hand, being the dominant group in organizational and societal life, might not depend *solely* on the co-authorship ties to produce research. They might have other resources such as research funding/grants and research equipment at their disposal. Hence their reliance on co-authorship ties (and consequently high centrality) for higher research performance might be less than that of female researchers. Hence co-authorship ties and consequently centrality might be more meaningful for female researchers as compared to male researches.

To sum up, academic women in Pakistan might use co-authorship network ties to *break the ice*. Therefore these exceptional women might be the ones who derive stronger benefits from co-authorship network ties and, hence, from their central position (degree, closeness and betweeness) in the network. Consequently we can propose the following hypothesis:

H2a Gender will moderate the relationship between degree centrality and research performance such that this relationship is stronger for female authors as compared to their male counterparts.

H2b Gender will moderate the relationship between closeness centrality and research performance such that this relationship is stronger for female authors as compared to their male counterparts.

H2c Gender will moderate the relationship between betweeness centrality and research performance such that this relationship is stronger for female authors as compared to their male counterparts.

Method

Network boundary and data

This particular study was carried out in Pakistan a 3rd world country. Since the inception of Higher Education Commission (HEC) in 2002, research output originating from Pakistan has significantly increased. HEC quotes:

"According to the Institute of Scientific Information (U.S.), the total number of publications appearing in the 8,000 leading journals indexed in the Web of Science arising out of Pakistan in 2005 was 1,259 articles, representing a 41 % increase over the past 2 years, and a 60 % increase since the establishment of the HEC in 2002."

Similarly there has been an increase in research output in following years as well with 1,759 articles in 2006, 2,494 in 2007, 3,640 in 2008, 4,143 in 2009 and 5007 in 2010 (See HEC annual reports: 2002–2010).

The first necessity was to identify and bound the population of authors in the co-authorship network. Bounding a social network is a critical step in network analysis (Hanneman and Riddle 2005). Social network analysts rarely use samples in their work. Most commonly, network analysts identify a population and conduct a census of that population. The boundaries are those imposed by the researcher or even created by the actors themselves. Social network studies, therefore often draw the boundaries around a population that is known, a priori, to be a network (Hanneman and Riddle 2005).

The network for this study consisted of all researchers employed/enrolled in universities and research institutes in Pakistan (and all their co-authors employed/enrolled in universities and research institutes in Pakistan as well) who have published in the field of Chemistry in ISI indexed journals from years 2002 to 2009. The year 2002 was chosen as a benchmark year because it is year in which HEC was established; Chemistry was chosen because it is considered as top research area (according to research output or number of publications) as described by HEC.

Bibliometric data for the study was collected by performing an ISI Web of Science (SCI) search from 2002 to 2009, keeping articles in the document type field and Pakistan in the address field. The search was refined to exclude all countries except Pakistan and further refined to include Chemistry and its sub-fields. This yielded a "domestic" co-authorship network (Kwon et al. 2012; Leydesdorff and Sun 2009) i.e. Pakistani authors co-authoring with Pakistani authors in Chemistry and its sub-fields only. The *raw* co-authorship network consisted of 1,814 articles with 3,008 authors.

Data cleaning

The *raw* bibliometric data cannot be used effectively for co-authorship networks analysis because of the problem of *author name disambiguation* (Smalheiser and Torvik 2009). Author name disambiguation problem can be characterized by homonymy and synonymy in bibliometric data (Guns et al. 2011).

Homonymy is quite common in bibliometric data and so was the case with the study's data set comprising of Pakistani names. For instance in the data set a last-name-plus-initial, Ahmad, S represented many different authors such as Ahmad, Safeer, Ahmad, Shujjat, Ahmad, Saeed, Ahmad, Shabir, Ahmad, Shakeel ect.

Synonymy was also common in our data. For instance a single author Bangash, Fazalullah Khan was present as Bangas, Fazlullah Khan (misspelled), Bangash, FK, Bangash, Fazalullah Khan (again misspelled) and Khan-Bangash, Fazalullah.

Hence in order to accurately study co-authorship networks the *author name disambiguation* problem had to be solved by cleaning the raw data (Gossart and Özman 2009).

The dataset provides records in both last-name-plus-initial and last-name-plus-full-first-name form (Strotmann et al. 2009). It was decided to adopt all possible measures to convert the names that were in last-name-plus-initial form to last-name-plus-full-first-name form to solve the problem of homonymy and synonymy (Strotmann et al. 2009). These measures include:

- Searching the articles on search engines such as Google and consequently in journals in which they are published and/or databases in which the journal is indexed;
- Searching the authors along with their institutional affiliation on search engines;
- Searching the email id's of authors on search engines;
- Checking the publications on personal WebPages and/or Curriculum vitas of authors.
- Searching the dataset itself for patterns of co-authorship based on the argument that authors can often be recognized on the basis of their co-authors (Kang et al. 2008).

Despite adopting all these measures, the dataset still contained ambiguous author names in last-name-plus-initial form which could not be converted in last-name-plus-full-first-name form. Those authors/papers were removed from the dataset. In addition some uniquely identifiable author names in last-name-plus-initial form were retained when a full name could not be obtained. Hence the final dataset consisted of 1,699 paper and 2,027 unique authors publishing in 151 journals. Full names of 1,845 unique authors were known and the rest (182 authors) are those whose names are in uniquely identifiable last-name-plus-initial form.

Measures

Dependent variable: research performance

We measured research performance for each author by weighting each publication of an author by the journal's 5 year impact factor.[3] The 5 year impact factor of each of the 151 journals in the study was accessed using ISI Web of Science (SCI) Journal Citation Reports (JCR). For each author, the weights by impact factors for each publication were added to create an aggregate research performance score. For example if an author publishes three papers, one in Photochemical & Photobiological sciences (5 year impact factor: 2.505), one in Journal of Chemical Society of Pakistan (5 year impact factor: 0.221) and one in Chemical & Pharmaceutical Bulletin (5 year impact factor: 1.621) the value for that author on variable of research performance will be 4.347 $[(1 \times 2.505 + 1 \times 0.221 + 1 \times 1.621)]$. Similarly another author publishing the same number of papers i.e. three, with two in Journal of the Chilean Chemical Society (5 year impact factor: 0.629) and one in Acta Chromatographica (5 year impact factor: 0.898) will have research performance value of 2.156 $[(2 \times 0.629) + (1 \times 0.898)]$.

Independent construct: centrality

The normalized versions of the three classical measures of centrality used in this study as independent variables were measured using social network analysis (SNA) techniques. The cleaned dataset was loaded in SCi^2 (Sci^2 Team 2009) in standard CSV format. The co-authorship network was then extracted using same software where nodes were authors and relationship between them were established when they co-author a publication. The co-authorship network was a binary un-directed (2,027 × 2,027) matrix. The extracted co-authorship matrix file was saved using Pajek.net extension and subsequently imported in UCINET VI (Borgatti et al. 2002) for further analysis. It was expected that the

[3] The 5-year journal Impact Factor is the average number of times articles from the journal published in the past five years have been cited in the JCR year.

co-authorship network will be a larger interconnected component with other smaller disconnected components (Yan and Ding 2009) and visualization showed it was indeed the case (1,782 authors). UCINET VI was used to extract the largest interconnected component (1,782 × 1,782 binary un-directed matrix). Normalized versions of centrality within the largest interconnected component were calculated using UCINET VI.

Moderating variable: gender

As informed earlier, full names of 1,845 authors were known hence gender of each researcher is assessed by simply looking at the full names of the researchers. Because all the author names are Pakistani names the gender of the authors was based on the cultural norms of gender based naming conventions. In addition, the dataset was verified for gender based naming conventions by two faculty members of the Literature and Languages and Gender Development Studies departments, respectively, of University of Balochistan, Pakistan. Gender is coded as binary variable (male = 1; female = 0).

Data analysis techniques

Correlation and regression analysis

As previously argued, conceptual overlaps between the three centrality measures were expected given they are measures of the same construct, i.e. centrality. Thus, these measures were also expected to be correlated (Bolland 1998; Valente and Foreman 1998; Faust 1997; Valente et al. 2008). In addition, according to our hypotheses, these measures were also expected to be correlated with the variable of research performance. The variables were also expected to be non-parametric in nature (Yan and Ding 2009). Hence, Spearman's correlation between centrality measures was used to test correlation amongst variables.

The indicator of research performance was a continuous variable with a highly skewed distribution. One of the options for modeling continuous variables with skewed distribution is to log-transform the variable. This option was valid because the variable did not contain zeros so there was no potential loss of data. UCINET's *node-level regression* was run to test the hypotheses, using the interaction terms (i.e. nrmdegreecentrality*X*gender, nrmclosenesscentrality*X*gender, nrmbetweenesscentrality*X*gender) to test the moderating effect of gender (variables were *z*-transformed prior to calculation of interaction terms). Node level regression computes basic linear regression statistics by OLS, and estimates standard errors and significance[4] using the random permutations (1,000 by default) method for constructing sampling distributions of R^2 and slope coefficients (Hanneman and Riddle 2005). This procedure was used because of the fact that observations might not be independent (since all the authors are part of the same network) and, hence to deal with autocorrelation. Variance inflation factor (VIF) values were also used to test for the issues of multicollinearity. The residual distribution was checked for normality assumptions (using histogram, skewness and kurtosis values, and normal P–P and Q–Q plots). In addition the Durbin–Watson statistic was used to test for independence/non-independence or autocorrelation of residuals.

[4] The *p* value for each statistic is calculated as the proportion of permutations that yields a statistic as extreme as the one initially produced.

Controlling for reverse causality (temporal analysis)

Understanding the fact that there might be ambiguity over the direction of causality in the model (does centrality lead to higher research performance or does higher research performance makes the author central?), the analysis was repeated, controlling for reverse causality using temporal analysis. The original CSV data file was split in 2 distinct time windows (time window 1: 2002–2005 and time window 2: 2006–2009) using SCi^2 (Sci2 Team 2009). Both time windows were compared to reveal common authors. This allowed testing, whether centrality (degree, closeness, betweeness) in time 1 (calculated using UCINET VI) leads to higher research performance in time 2 and whether the relationship is stronger for females as compared to males. Spearman correlation and node level regression was again run with all the statistics described in previous sub-section.

Results and discussion

SNA

As discussed earlier, SNA showed that the network indeed had one largest interconnected component (1,782 authors) out of which we knew full names (and hence gender) of 1,628 authors. The rest (246 authors) were present in various disconnected clusters. Hence 1,628 authors (439 female authors, 1,189 male authors) were finally selected for hypotheses testing.

Correlation and regression analysis

Table 1 reports means, standard deviation and spearman correlation statistic for the relationship between each study variable at large component level. These findings show a positive and significant correlation between variables at the 0.01 level.[5]

Result of the node-level multiple regression is reported in Table 2. All the VIF values were below 5 indicating no problem of multicollinearity. The residual distribution was normal (histogram looked normal and normal P–P and Q–Q plot were almost a straight line). In addition skewness and kurtosis were near zero (skewness = 0.076 and kurtosis = −0.357). The Durbin–Watson statistic was near 2 (1.950) indicating independence (almost no autocorrelation) of residuals. The regression model included the independent variables gender and interaction terms ($R^2 = 0.306$). The coefficients for degree and closeness centrality were positive and significant ($\beta_1 = 0.568$, $p < 0.01$; $\beta_2 = 0.139$, $p < 0.01$). Hence we find support for hypotheses H1a and H1b. Surprisingly though, the coefficient for betweeness was negative and significant ($\beta_3 = -0.130$, $p < 0.01$) finding no support for H1c. Coefficient for gender was insignificant ($\beta_4 = 0.023$, $p > 0.10$). Coefficients for interactions term were negative and significant only for degree centrality indicating the relationship between degree centrality and research performance to be stronger for female authors as compared to male authors ($\beta_5 = -0.128$, $p < 0.05$). Hence H2a was supported. The coefficients for interaction terms for closeness centrality and betweeness centrality were positive but insignificant (Model 3: $\beta_6 = 0.060$, $p > 0.05$; $\beta_7 = 0.053$, $p > 0.10$) hence H2b and H2c were not supported.

[5] Distributions for degree centrality and betweeness centrality followed highly skewed curves. Distribution for closeness centrality followed normal curve.

Table 1 Means, standard deviations and Spearman correlations (non-temporal analysis)

	Mean	SD	1	2	3
1. Research performance	0.251	1.359	–		
2. Degree centrality	0.432	0.573	0.558[a]	–	
3. Closeness centrality	21.326	3.708	0.384[a]	0.562[a]	–
4. Betweeness centrality	0.236	0.853	0.563[a]	0.751[a]	0.382[a]

[a] Correlation is significant at the 0.01 level (2-tailed)

Table 2 Results of node-level regression for research performance (non-temporal analysis)

	Un-standardized coefficients	Standardized coefficients	VIF
Intercept	0.253	0.000	
Degree centrality	0.773**	0.568**	3.472
Closeness centrality	0.189**	0.139**	1.347
Betweeness centrality	−0.177**	−0.130**	3.199
Gender	0.032	0.023	1.019
Degree centrality × gender	−0.194**	−0.128**	2.758
Closeness centrality × gender	0.081	0.060	1.341
Betweeness centrality × gender	0.088	0.053	2.523
R^2	0.306**		
Adjusted R^2	0.302**		

p values based on permutation tests (1,000 repetitions) ** $p < 0.01$

Figure 1 presents the interaction plots to visually inspect the moderator effect of gender for degree centrality and research performance. Figure 1 graphs the un-standardized predicted values for female and male authors. The figure clearly depict that the relationship is stronger for female authors as compared to male authors.

Controlling for reverse causality (temporal analysis)

SNA

Results of SNA in time window 1 (2002–2005) showed presence of 956 authors, 756 of whom were in largest connected component. As discussed before, both time windows were compared to reveal 442 authors, out of which 395 were in largest connected component of time window 1 (2002–2005). Fortunately full names of all 395 authors were known so 395 authors (95 female authors, 300 male authors) were finally selected for hypotheses testing across time. Research performance of 395 authors was calculated in time window 2 (2006–2009).

Correlation and regression analysis

Table 3 reports means, standard deviation and spearman correlation. The relationship between variables showed a positive and significant correlation at the 0.01 level. The

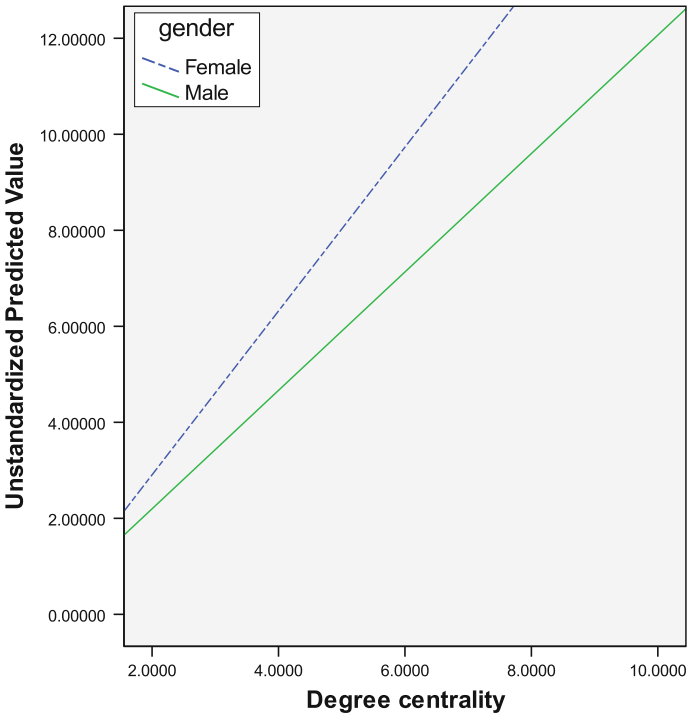

Fig. 1 Interaction plot for moderating impact of gender (non-temporal analysis)

Table 3 Means, standard deviations and Spearman correlations (temporal analysis)

	Mean	SD	1	2	3
1. Research performance	0.673	1.326			
2. Degree centrality	1.147	1.358	0.309[a]		
3. Closeness centrality	18.214	3.814	0.301[a]	0.477[a]	
4. Betweeness centrality	1.101	2.780	0.215[a]	0.764[a]	0.275[a]

[a] Correlation is significant at the 0.01 level (2-tailed)

correlation between research performance and centrality measures is weaker when compared to the initial or non-temporal analysis.

Result of the node-level multiple regression is reported in Table 4. All the VIF values were below 5 indicating no problem of multicollinearity. The residual distribution was normal (histogram looked normal and normal P–P and Q–Q plot were almost a straight line). In addition skewness and kurtosis were near zero (skewness = 0.138 and kurtosis = −0.296). The Durbin–Watson statistic was near 2 (1.845) indicating independence (almost no autocorrelation) of residuals. The regression model included the independent variables gender and interaction terms ($R^2 = 0.204$). The coefficients for degree and closeness centrality were positive and significant ($\beta_1 = 0.348$, $p < 0.01$; $\beta_2 = 0.175$, $p < 0.05$). Hence we find support for hypotheses H1a and H1b. Coefficient for betweeness was negative and insignificant ($\beta_3 = -0.012$, $p > 0.10$) finding no support for H1c.

Table 4 Results of node-level regression for research performance (temporal analysis)

	Un-standardized coefficients	Standardized coefficients	VIF
Intercept	0.674	0.000	
Degree centrality	0.462**	0.348**	2.099
Closeness centrality	0.232*	0.175*	1.248
Betweeness centrality	−0.015	−0.012	1.706
Gender	−0.030	−0.021	1.012
Degree centrality × gender	−0.155	−0.094	1.833
Closeness centrality × gender	−0.146***	−0.104***	1.225
Betweeness centrality × gender	0.152	0.102	1.439
R^2	0.204**		
Adjusted R^2	0.188**		

p values based on permutation tests (1,000 repetitions) *** $p < 0.10$; ** $p < 0.01$; * $p < 0.05$

Coefficient for gender was insignificant ($\beta_4 = 0.021$, $p > 0.10$). In this temporal analysis, coefficient for interactions term were negative and significant only for closeness centrality indicating that the relationship between closeness centrality and research performance is stronger for female authors as compared to male authors ($\beta_5 = -0.104$, $p < 0.10$). Hence H2b was supported. The coefficients for the interaction terms for degree and betweeness centrality were negative but insignificant and positive but insignificant respectively ($\beta_6 = -0.094$, $p > 0.10$; $\beta_7 = 0.102$, $p > 0.10$). Hence H2a and H2c were not supported.

Figure 2 present the interaction plots to visually inspect the moderator effect of gender for closeness centrality and research performance. Figure 2 graphs the un-standardized predicted values for female and male authors. The figure depict that the relationship is stronger for female authors as compared to male authors.

The initial or non-temporal analysis confirmed the association of centrality and research performance and the temporal analysis confirmed the direction of causality. Pakistani authors publishing in chemistry, embedded in their domestic co-authorship network, seem to benefit from their direct and distinct co-authorship ties (degree centrality) and quick flow of knowledge and skills by the virtue of having low average distance to other authors (closeness centrality) and consequently have higher research performance, as measured by the aggregate ISI impact factors of their publications. The surprising finding in both non-temporal and temporal analysis is related to the betweeness centrality of authors. In both analyses, betweeness centrality was negatively associated with research performance, (although insignificant in the temporal analysis). The reason might be the costs associated with diverse and non-redundant knowledge access. A consequence of a having high volume of more diverse or non-redundant knowledge access might be that it consumes time and resources that then cannot be allocated for absorbing and integrating the obtained knowledge and skills (Gilsing et al. 2008). A second cost may result for authors with high betweenness as they may have more non-redundant ties, which could result in a knowledge drift such that an author's knowledge base might change continuously in different and unrelated directions, making the accessed diverse knowledge difficult to absorb and integrate (Fleming and Sorenson 2001; Ahuja and Katila 2004). While similar costs can also be associated with degree centrality (direct co-authorship relationships requiring time, energy and attention to establish), and closeness centrality (author's being aware of too many conflicting view points within the network), yet Pakistani authors in their domestic co-authorship network seem to derive good amount of benefits from central positions based

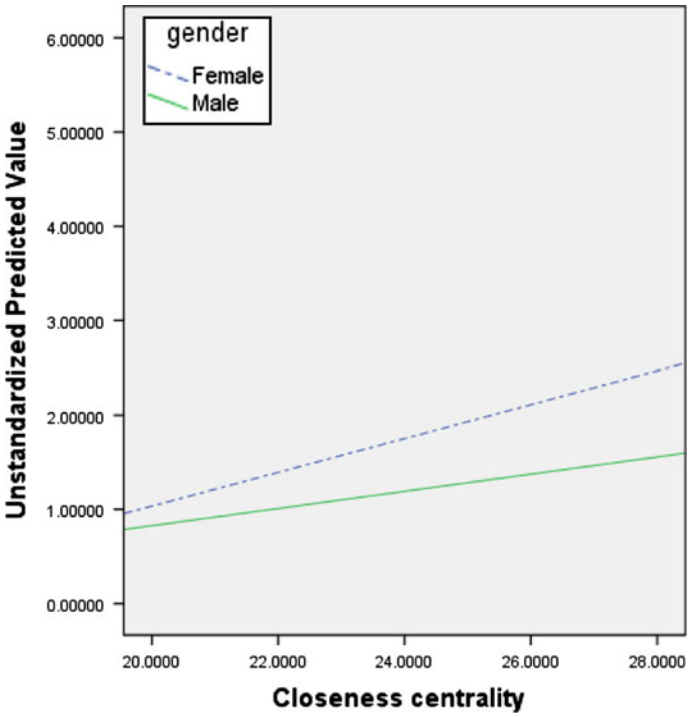

Fig. 2 Interaction plot for moderating impact of gender (temporal analysis)

on degree and closeness. Perhaps degree centrality was beneficial given the context of this study. This study focused only on the domestic co-authorship network of Pakistani authors publishing in Chemistry (faculty/researchers employed and students enrolled in Pakistani Universities/research institutes), in which, Pakistani faculty members can establish co-authorship relations with the students enrolled under their supervision and other faculty members of same or different institutions within the country, relatively easily. In addition, perhaps closeness was beneficial given that the network might not offer too many conflicting view points because all authors belong to the same country and are publishing in same field. Yet, on the other hand, the focus on this domestic co-authorship network (faculty/researchers employed and students enrolled in Pakistani Universities/research institutes) might also explain the negative and/or insignificant impact of betweeness on research performance given that research faculty may be the ones who broker the connections between students. It is understandable that students might not possess ample knowledge and skills which faculty can utilize as brokers. This scenario, along with the costs associated with brokerage (mentioned above) might not be a fruitful proposition for faculty.

Findings related to the moderating impact of gender indicated a stronger relationship between degree centrality and research performance for Pakistani female authors (non-temporal analysis) implying that they derive more benefits from direct and distinct co-authorship ties as compared to Pakistani male authors. Temporal analysis revealed a stronger relationship between closeness centrality and research performance for Pakistani female authors. This specifies that, during the initial phase of establishment of the

co-authorship network (i.e. time window 1: 2002–2005), female authors seem to utilize the quick flow of knowledge and skills (by the virtue of having low average distance to other authors) better than male authors and, hence, in a later time (i.e. time window 2: 2006–2009) have higher research performance as compared to men. With the establishment of the co-authorship network over time (2002–2009), female authors seem to utilize the knowledge and skills from direct co-authorship (degree centrality) ties, better than male authors, to publish in more or higher impact factor journals. These findings strengthen our proposed argument that awareness of female authors about their co-authorship network might be shaping the structure of co-authorship network in their favor. Therefore during the initial establishment of co-authorship network, when there might be few direct ties (degree centrality) to be utilized, they make use of overall structure of the network utilizing knowledge and skills from the direct as well as indirect ties (closeness centrality). However, with the establishment of the co-authorship network over time, and with the availability of more direct and distant co-authorship ties, they shift their focus towards utilization of knowledge and skills from direct ties (degree centrality). They might be aware that now that the co-authorship network is established containing many authors, utilizing knowledge and skills from indirect ties (closeness) centrality might be costly.

Conclusions and recommendations

This study offers several academic implications for developing country researchers (particularly for researchers in Pakistan) as well as recommendations for future research. This research examined the domestic co-authorship network of Pakistani researchers publishing in Chemistry and its sub-fields (using SNA) and the association of their network centrality (degree, closeness and betweeness) and research performance (aggregate of publications weighted by 5 year impact factors) as well as how a particular individual characteristic, gender moderates this relationship. Both Non-temporal and temporal analyses were used to test the hypotheses. Temporal analysis in particular, leads to an important implication of potential causality. It implies how author network centrality during an earlier time period positively predicts research performance during a later time period. In addition, use of an individual characteristic, gender as a moderating variable has important implications for social network research. It implies that social side of performance might not be a straight forward proposition and individuals possessing certain attributes might benefit more/less from their social networks. Access to CV's of research faculty might inform the addition of more *individual* as well as *work related characteristics* which can act as moderating variables (such as rank, age, affiliation etc.).

The findings indicate the positive relationship of degree (direct ties) and closeness centrality (quick flow of knowledge) on research performance of these Pakistani authors, with a stronger relationship for female authors. Pakistani researchers could use the findings of this study to devise co-authorship strategies to improve their research performance. The findings could be heartening for female faculty and female students employed and enrolled in Pakistani institutes. Policy makers and research institutes in Pakistan might be encouraged to devise strategies to motivate and facilitate female academics into publish their research in impact factor journals. Female authors in the country could use co-authorship ties to surpass barriers faced in male dominated institutes. Future research could replicate these analyses in other academic domains, such as Engineering, Humanities and Social sciences etc. and other developing countries.

The surprising finding of the negative relationship of betweenness centrality (brokerage and control) on research performance might be tested further by dropping out students and studying the network of only the research faculty. Moreover, expanding the domestic Pakistani co-authorship network to an international co-authorship network of Pakistani researchers might serve the above purpose as well. In addition, it might provide more insightful findings related to the moderating effect of gender and address the question of "do Pakistani female authors derive more benefits than male authors from their central position in the international co-authorship network?"

References

Abbasi, A., Altmann, J., & Hossain, L. (2011). Identifying the effects of co-authorship networks on the performance of scholars: A correlation and regression analysis of performance measures and social network analysis measures. *Journal of Informetrics, 5*(4), 594–607.

Ahuja, G. (2000). Collaboration networks, structural holes, and innovation: A longitudinal study. *Administrative Science Quarterly, 45*(3), 425–455.

Ahuja, G., & Katila, R. (2004). Where do resources come from? The role of idiosyncratic situations. *Strategic Management Journal, 25*(8/9), 887–907.

Albrecht, S. L. (1983). Informal interaction patterns of professional women. In J. R. Gordon (Ed.), *A diagnostic approach to organizational behavior* (pp. 287–290). Boston: Allyn & Bacon.

Arora, A., & Gambardella, Ae. (1990). Complementarity and external linkages: The strategies of the large firms in biotechnology. *Journal of Industrial Economics, 34*(4), 361–379.

Avital, M., & Collopy, F. (2001). Assessing research performance: Implications for selection and motivation. *Sprouts: working papers on information environments, systems and organizations*. http://sprouts.aisnet.org/1-14. Accessed 10 March 2010.

Berg, S., Duncan, J., & Friedman, P. (1982). Joint venture strategies and corporate innovation. Cambridge, MA: Oelgeschiager, Gunn & Hain.

Bhardwaj, A., Qureshi, I., & Lee, S. H. (2008). A study of race/ethnicity as a moderator of the relationship between social capital and satisfaction. *Academy of management annual meeting*. Anaheim, California.

Bolland, J. M. (1998). Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. *Social Networks, 10*, 233–253.

Borgatti, S. P. (1995). Centrality and AIDS. *Connections, 18*(1), 112–114.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *UCINET for windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.

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. (1992). *Structural holes—The social structure of competition*. Cambridge, MA: Harvard University Press.

Burt, R. S. (1998). The gender of social capital. *Rationality and Society, 10*, 15–46.

Burt, R. S. (2005). *Brokerage and closure: The social capital of structural holes*. Oxford: Oxford University Press.

Cartwright, V. A., & McGhee, C. N. (2005). Ophthalmology and vision science research part 1: Understanding and using journal impact factors and citation indices. *Journal of Cataract and Refractive Surgery, 31*, 1999–2007.

Cheek, J., Garnham, B., & Quan, J. (2006). What's in a number? Issues in providing evidence of impact and quality of research(ers). *Qualitative Health Research, 16*, 423–435.

Coleman, J. S. (1990). *Foundations of social theory*. Cambridge, MA: Belknap Press of Harvard University Press.

C-Zurián, J., Alcaide, G. G., J-Zurián, F. J., Benavent, & Miguel-Dasit, A. (2007). Coauthorship Networks and Institutional Collaboration in Revista Española de CardiologÍa Publications. *Revista Espanola de Cardiologia, 60*(2), 117–130.

Department of Economic and Social Affairs. (2010). *The World's Women*. New York: United Nations.

Eaton, J. P., Ward, J. C., Kumar, A., & Peter, H. R. (1999). Structural analysis of co-author relationships and author productivity in selected outlets for consumer behavior research. *Journal of Consumer Psychology, 8*(1), 39–59.
Fairhurst, G. T., & Snavely, B. K. (1983). A test of the social isolation of male tokens. *Academy of Management Journal, 26*, 353–361.
Faust, K. (1997). Centrality in affiliation networks. *Social Networks, 19*, 157–191.
Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy, 30*, 1019–1039.
Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35–41.
Freeman, L. C. (1979). Centrality in social networks. Conceptual clarification. *Social Networks, 1*, 215–239.
Gilsing, V., Nooteboomb, B., Vanhaverbekec, W., Duystersd, G., & Oorda, Av. (2008). Network embeddedness and the exploration of novel technologies: Technological distance, betweenness centrality and density. *Research Policy, 37*, 1717–1731.
Gossart, C., & Özman, M. (2009). Co-authorship networks in social sciences: The case of Turkey. *Scientometrics, 78*(2), 323–345.
Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*, 481–510.
Gulati, R., & Martin, G. (1999). Where do interorganizational networks come from? *American Journal of Sociology, 104*(5), 473–506.
Guns, R., Liu, Y. X., & Mahbuba, D. (2011). Q-measures and betweenness centrality in a collaboration network: A case study of the field of informetrics. *Scientometrics, 87*(1), 133–147.
Hanneman, R. A., & Riddle, M. (2005). *Introduction to social network methods*. Riverside CA: Analytictech.com.
Hendrick, S. S. (1981). Why women don't succeed. *National Business Employment Weekly, 40*, 9–11.
Higher Education Commission, Pakistan. http://beta.hec.gov.pk/InFocus/Pages/HEC_Report_for_2002-2008.aspx. Accessed 17 March 2012.
Higher Education Commission, Pakistan. http://beta.hec.gov.pk/Pages/HECMain.aspx. Accessed 17 March 2012.
Ibarra, H. (1993). Network centrality, power and innovation involvement: Determinants of technical and administrative roles. *Academy of Management Journal, 36*, 471–501.
Ibarra, H. (1997). Paving an alternative route, gender differences in managerial networks. *Social Psychology Quarterly, 60*, 91–102.
Jansen, D., Gortz, R., & Heidler, R. (2010). Knowledge production and the structure of collaboration networks in two scientific fields. *Scientometrics, 83*(1), 219–241.
Kang, H., Getoor, L., Shneiderman, B., Bilgic, M., & Licamele, L. (2008). Interactive entity resolution in relational data: A visual analytic tool and its evaluation. *IEEE Transactions on Visualization and Computer Graphics, 14*(5), 999–1014.
Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
Kay, F., & Hagan, J. (1999). Cultivating clients in the competition for partnership: Gender and the organizational restructuring of law firms in the 1990s. *Law and Society Review, 33*, 517–555.
Kram, K. E. (1988). *Mentoring at work: Developmental relationships in organizational life*. New York: University Press of America.
Kretschmer, H., & Aguillo, I. F. (2005). New indicators for gender studies in web networks. *Information Processing and Management, 41*(6), 1481–1494.
Kwon, K. S., Park, H. W., So, M., & Leydesdorff, L. (2012). Has globalization strengthened South Korea's national research system? National and international dynamics of the triple helix of scientific co-authorship relationships in South Korea. *Scientometrics, 90*(1), 163–176.
Leydesdorff, L., & Sun, Y. (2009). National and international dimensions of the triple helix in Japan: University–industry–government versus international co-authorship relations. *Journal of the American Society for Information Science and Technology, 60*(4), 778–788.
Liao, C. H. (2011). How to improve research quality? Examining the impacts of collaboration intensity and member diversity in collaboration networks. *Scientometrics, 86*(3), 741–761.
Marsden, P. V. (2002). Egocentric and sociometric measures of network centrality. *Social Networks, 24*, 407–422.
Mcfadyen, A. M., & Cannella, J. A. (2004). Social capital and knowledge creation: Diminishing returns of the number and strength of exchange relationships. *Academy of Management Journal, 47*(5), 735–746.
McFadyen, A. M., Semadeni, M., & Cannella, A. A, Jr. (2009). The value of strong ties to disconnected others: Examining knowledge creation in biomedicine. *Organization Science, 20*(3), 552–564.
Meho, L. I. (2007). The rise and rise of citation analysis. *Physics World, 20*, 32–36.

Mehra, A., Kilduff, M., & Brass, D. (2001). The social networks of high and low-self monitors: Implications for workplace performance. *Administrative Science Quarterly, 46*, 121–146.
Nagpaul, P. S. (2002). Visualizing cooperation networks of elite institutions in India. *Scientometrics, 54*(2), 213–228.
Nahapiet, J., & Ghoshal, S. (1998). Social capital, intellectual capital, and the organizational advantage. *Academy of Management Review, 23*, 242–266.
Newman, M. E. (2001a). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, *64*(1), 016132.
Newman, M. E. (2001b). The structure of scientific collaboration networks. *PNAS, 98*, 404–409.
Newman, M. E. (2004). Co authorship networks and patterns of scientific collaboration. *PNAS, 101*, 5200–5205.
Newman, M. E. (2010). *Networks: An introduction*. Oxford University Press.
Oh, W., Choi, Jn, & Kim, K. (2005). Co-authorship dynamics and knowledge capital: The patterns of cross-disciplinary collaboration in information systems research. *Journal of Management Information Systems, 22*(3), 265–292.
Perry-Smith, J. E., & Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *The Academy of Management Review, 28*(1), 89–106.
Pike, T. W. (2010). Collaboration networks and scientific impact among behavioral ecologists. *Behavioral Ecology, 21*(2), 431–435.
Ragins, B. R., & Sundstrom, E. (1989). Gender and power in organizations: A longitudinal perspective. *Psychological Bulletin, 105*, 51–88.
Richardson, G. B. (1972). The Organisation of Industry. *The Economic Journal, 82*(327), 883–896.
Sci2 Team. (2009). Science of Science (Sci2) Tool. *Indiana University and SciTech Strategies*. http://sci2.cns.iu.edu. Accessed 5 May 2011.
Scott, J. (1991). *Social network analysis: A handbook*. Sage.
Seron, C., & Ferris, K. (1995). Negotiating professionalism: The gendered social capital of flexible time. *Work and Occupations, 22*, 22–47.
Smalheiser, N. R., & Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology, 43*, 287–313.
Sparrowe, T., Liden, R., Robert, G., Wayne, J. S., & Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of Management Journal, 44*(2), 316–325.
Strotmann, A., Zhao, D., & Bubela, T. (2009). Author name disambiguation for collaboration network analysis and visualization. *Proceedings of the American Society for Information Science and Technology, 46*(1), 1–20.
Tharenou, P. (1999). Gender differences in advancing to the top. *International Journal of Management Reviews, 1*, 111–132.
Tsai, W. (2001). Knowledge transfer in intraorganizational networks: Effects of network position and absorptive capacity on business unit innovation and performance. *Academy of Management Journal, 44*(5), 996–1004.
Tsai, W., & Ghoshal, S. (1998). Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal, 41*, 464–476.
Valente, T. W., & Foreman, R. (1998). Integration and radiality: Measuring the extent of an individual's connectedness and reachability in a network. *Social Networks, 20*, 89–109.
Valente, T. W., Loronges, K., Lakon, C., & Costenbader, E. (2008). How correlated are network centrality measures? *Connections, 28*(1), 16–26.
Wagner, C. S. (2006). International collaboration in science and technology: Promises and pitfalls. In L. Box & R. Engelhard (Eds.), *Science and technology policy for development: Dialogues at the interface*. London: Anthem Press.
Wasserman, S., & Faust, K. (1994). *Social networks analysis: Methods and applications*. Cambridge: Cambridge University Press.
Yan, E., & Ding, Y. (2009). Applying centrality measures to impact analysis: A coauthorship network analysis. *Journal of the American Society for Information Science and Technology, 60*(10), 2107–2118.
Yang, K.-s. (2007). Firms' network position, industry hierarchy position and innovation and an additional examination on structural equivalent block models and between-sector position. *Academy of management annual meeting*. Philadelphia, PA.
Yousefi-Nooraie, R., Akbari-Kamrani, M., Hanneman, R. A., & Etemadi, A. (2008). Association between co-authorship network and scientific productivity and impact indicators in academic medical research centers: A case study in Iran. *Health Research Policy and Systems*, *6*(9). doi:10.1186/1478-4505-6-9.
Zheng, W. (2008). A social capital perspective of innovation from individuals to nations: Where is empirical literature directing us? Volume 10 Issue 4. *International Journal of Management Reviews, 10*(4), 1–39.