Journal of Informetrics 7 (2013) 683–692

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

Research impact and scholars' geographical diversity

Alireza Abbasi[a,*], Ali Jaafari[b]

[a] School of Engineering and IT, University of New South Wales, Canberra, ACT 2610, Australia
[b] Graduate School of Business and Project Management, Asia Pacific International College, Sydney, NSW 2008, Australia

ARTICLE INFO

Article history:
Received 16 July 2011
Received in revised form 12 March 2013
Accepted 15 April 2013
Available online 22 May 2013

Keywords:
Scientific collaboration network
Geographical diversity
Internal and external collaboration
Research impact
Scholars' performance

ABSTRACT

In recent years there has been a sharp increase in collaborations among scholars and there are studies on the effects of scientific collaboration on scholars' performance. This study examines the hypothesis that geographically diverse scientific collaboration is associated with research impact. Here, the approach is differentiated from other studies by: (a) focusing on publications rather than researchers or institutes; (b) considering the geographical diversity of authors of each publication; (c) considering the average number of citations a publication receives per year (time-based normalization of citations) as a surrogate for its impact; and (d) not focusing on a specific country (developed or developing) or region. Analysis of the collected bibliometric data shows that a publication impact is significantly and positively associated with all related geographical collaboration indicators. But publication impact has a stronger association with the numbers of external collaborations at department and institution levels (inter-departmental and inter-institutional collaborations) compared to internal collaborations. Conversely, national collaboration correlates better with impact than international collaboration.

1. Introduction

Most of the research problems and their solutions are borderless and globalization is an inevitable phenomenon in research (Barjak & Robinson, 2008). The rapid growth of scientific international collaboration over the past decades has attracted the attention of many researchers in different disciplines. They have attempted not only to investigate the increment rate and its pattern but also to argue its reasons and effect on the science and industry. Naturally, it is acceptable that diverse knowledge and skills from different backgrounds or research cultures is often considered beneficial for science (Barjak & Robinson, 2008) but it might have its obstacles such as language barriers, availability and accessibility to the resources for the whole group in an appropriate time.

Since individuals' resources and capacity (including time) are limited, their interactions with others are indispensable to perform large tasks efficiently. Scholars are no exception to this trend; and there is a need for cooperation (collaboration) to create new knowledge (Demsetz, 1991) through sharing their knowledge and skills to perform large scale research projects. Scientific research collaboration is a social process, "there are as many reasons for researchers to collaborate as there are reasons for people to communicate" (Katz, 1994, p. 32). For instance, many collaborative activities result in publications of joined work, dissemination of research results, and creation of new ideas (Borgman & Furner, 2002). In addition, public and private research funding agencies are requiring interdisciplinary, international, and inter-institutional collaborations (Sonnenwald, 2007).

* Corresponding author. Tel.: +61 405676636.
E-mail addresses: alireza.abbasi@unsw.edu.au, a.abbasi@unsw.edu.au (A. Abbasi).

1751-1577/$ – see front matter
http://dx.doi.org/10.1016/j.joi.2013.04.004

All too often, many scientific outputs are a result of group work or cooperation; and publications are one of the most visible surrogates reflecting output of an academic research project. Publications are used for evaluating and comparing publishing productivity of scholars which is needed for faculty recruitment, government and industry funding and so on. Consequently, publications are the most widely used tool not only for assessing the scholars and their institutes' productivity or performance (Abbasi, Hossain, & Owen, 2012) but also for investigating their collaboration activities.

Although the effects of the scientific research collaboration on the performance are well understood and are promoted through governmental science policy makers especially in international level but there is not enough research on the effect that the geographically diverse collaborators may have on the publications quality. It is often assumed that geographical proximity is required for group work and it facilitates collaboration due to the need for face-to-face interaction (Ponds, Van Oort, & Frenken, 2007). In one of the rare studies which scientifically examined the matter, Barjak and Robinson (2008) found that having a strong domestic base (local collaboration) but collaborating actively enough outside the country, "to ensure a moderate amount of external involvement in the team", is required for enriching the collaboration and leading to better research output. Ponds et al. (2007) claim that "geographical proximity can only have an indirect role, and is neither a prerequisite nor sufficient for successful collaboration" (p. 424).

In order to investigate further the relationship between the collaborators' geographical location and the quality of collaboration output, this study aims to answer the following research questions: Do collaborations among scholars lead to a higher impact of a research? If so, do internal collaborations (considering different levels) have the same effect as external collaborations on the research impact?

To do so, publication data of top 10 journals of the "Information Science & Library Science" were selected and extracted from Scopus. Focusing on the publications rather than authors (in co-authorship analysis), the frequency of intra- and inter-departmental, intuitional and national collaborations of each publication's authors was recorded. On the other hand, publications' citation counts were normalized, as a surrogate for publications' impact (quality). A correlation test was used to examine how each of these different levels of internal and external collaboration associate with the publication's impact.

The next section reviews the literature on research collaboration, research impact and collaboration's geographical diversity. Section 3, presents the data source and collection method, measures and methods of determining publication's collaboration measures (i.e., number of internal and external collaborations in department, institute and country levels) and publication's impact measure (i.e., normalized citations count). Besides, the research model and the associated hypothesis are outlined. The following section shows the statistics of the extracted database and the result of correlation test between the independent variables (i.e., the collaboration types) and the dependent variable (i.e., publication impact). Finally, the paper concludes by discussing the research limitations and highlighting the contributions.

2. Literature review

2.1. Research collaboration

Sharing resources such as knowledge, skills and tools to conduct large scale projects is common. Individual researchers often come along to work with each other by sharing their resources to achieve a common goal. Based on the individuals' level of involvement on the project, they might collaborate by working together almost equally or just cooperate by receiving some support from one or more parties.

In academia, the process of sharing resources to conduct research project called scientific collaboration even though sometimes it is actually cooperation. The cooperation usually leads to at least one output in terms of a publication or even a product or a service. Katz (1994) believes that scholars need to collaborate in much the same way as ordinary people needing to communicate. Katz (1994) has summarized many factors for promoting research collaboration. The main reasons are: "changing patterns of funding"; "popularity, visibility and recognition"; the need for specific tools and instruments; "the need to gain experience, train researchers"; enriching of ideas and techniques; and decreasing "spatial distance" (p. 2).

Many factors encourage the scientific collaboration; among these generating more output (and often with better quality) has been considered as the most important one. For instance, Melin (2000) found that sharing knowledge and skills generates more knowledge that leads to new ideas, better performance, and academic productivity. In his study, 38% of the scholars responding to his questionnaire chose 'increased knowledge' as the main benefit generated from the co-authorship collaboration. Also, many other studies, such as Abbasi, Chung, and Hossain (2012), Abbasi, Hossain, et al. (2012), Ahuja, Carley, and Galletta (1997), Barjak and Robinson (2008), Cross and Cummings (2004) and Yan and Ding (2009), found significant support for the positive association between scholars' number of collaborations and their performance. Therefore, the creation of new scientific knowledge, including new research questions, new research proposals, new theories, and new publications are considered as important results of the scientific collaboration (Stokols, Harvey, Gress, Fuqua, & Phillips, 2005).

Due to the necessity of collaboration as a policy for countries to keep pace with scientific progress (van Raan, 2004) many governments have been interested in enhancing the level of international collaborations (Katz & Martin, 1997; Melin & Persson, 1996) and research collaboration has become a major issue in the conduct of science (Melin & Persson, 1996). For instance, there have been policies to improve collaborations among individual researchers and between science and technology and numerous initiatives by the European Commission and national governments have been launched with the aim of developing collaboration among researchers (Barjak & Robinson, 2008; Katz & Hicks, 1997).

Scientometrics investigates patterns of collaboration using quantitative methods (Sonnenwald, 2007) through quantifying scholars' collaboration activity. In order to investigate collaboration, proper indicators are required to reveal different types of collaborations (Melin & Persson, 1996). Although there are different types of collaboration activities among scientists, including co-authorship, co-participation, co-supervision and other related data (Nankani, Simoff, Denize, & Young, 2009) but co-authorship collaborations among researchers in the community has been recognized as the most important type (Laudel, 2002) and the most visible and accessible type of collaboration (Jiang, 2008). Publications' collaboration activity is often measured through simple metrics such as the total number of authors or the number of internal and external authors.

2.2. Research impact

The number of citations a publication receives indicates not only that the work has passed the difficult process of acceptance but also it has been found relevant to someone else's work (Gerrity & McKenzie, 1978). Thus, citations qualify the quantity of publications and this makes it a good indicator to rank publications' value (Lehmann, Jackson, & Lautrup, 2006). Although there were debates about the real outcome of citations (Stigler & Friedland, 1975) but as even critical cites often indicate that the publication has contributed to a 'major intellectual debate' (Gerrity & McKenzie, 1978), it can be considered as a proxy for quality (Laband, 1985). Consequently, citations are used as a basis for more complicated measures and as a surrogate for performance of researchers, institutes and journals (e.g., researchers and communities' *h*-index and its variants, journals' impact factor). In summary, citation count is the most well-known and widely used metric to quantify publications' impact.

In order to find the way in which collaboration changes the impact of a research publication, several studies have explored how the number of citations received, average citations, and advanced citation-based performance measures (e.g., *h*-index, *g*-index) (Abbasi, 2013; Abbasi, Altmann, & Hossain, 2011; Abbasi, Chung, et al., 2012; Katz & Hicks, 1997; Narin, 1991; Yan & Ding, 2009; Yan, Ding, & Zhu, 2010). These vary with different types of collaboration.

As Sonnenwald (2007) pointed out, Avkiran's (1997) recommended to decision-makers to hesitate "in interpreting collaborative research as a definitive sign of ability to produce better research". Also, Duque et al. (2005) have found that collaboration is not associated with an increase in scientific publications in the developing countries of Ghana, Kenya, and the state of Kerala of India. Lee and Bozeman (2005) supported the positive association between the number of US scientific collaborations and the number of publications produced. In addition, van Raan (2004) and the Leiden group (plus many others) have also investigated the collaboration issue and found that there is a definitive correlation between (international) collaboration and impact at higher aggregation levels (institutional/country levels). Therefore, there is an open question regarding the correlation between research collaboration types (internal or external) and research impact (scientific performance).

2.3. Collaboration geographical diversity

Since diversity of members facilitates the integration of expertise, contribute to the successful project implementation and accelerate cycle time for new product development (Cummings, 2004; Eisenhardt & Tabrizi, 1995; Griffin & Hauser, 1992; Pinto, Pinto, & Prescott, 1993), researchers with different skills, experience and knowledge must work together in a cohesive group (McFadyen, Semadeni, & Cannella, 2009). This is in addition to basic shared understanding of each other's knowledge. Differentiating the interaction and collaboration internally (among members of a work group) and externally, Monge, Rothman, Eisenberg, Miller, and Kirste (1985) found that while internal and external knowledge sharing has direct implications on performance, the latter is more valuable when individuals in the workgroup are more structurally diverse. Also, Cummings and Cross (2003) have found that individuals in workgroups are more likely to perform better if they exchange knowledge externally with members in their professional network who are structurally diverse.

With rapid growth of international scientific collaboration, there have been many studies on the patterns of global collaboration (Georghiou, 1998; Glänzel, Schubert, & Czerwon, 1999; Gomez, Teresa Fernández, & Sebastian, 1999). For instance, Katz (1994), using bibliometric data focusing on university–university collaborations, examined the relationship between researchers' geographical locations and the amount of collaborations and found that the amount of collaboration among universities decreases in regards to their (geographical) distance. Also, Hoekman, Frenken, and Van Oort (2009) analyzed European inter-regional research collaboration in order to assess the effects of geographical and institutional distance on research collaboration. They found that distance has a significant negative effect on the chance and intensity of collaboration. In other words, geographically close regions collaborate more than regions spaced apart.

In one of the rare studies on researchers' level rather than country or regional levels, Barjak and Robinson (2008) examined the extent to which different countries of origin and cultural diversity of academic research teams are related to their research performance. They found that "the most successful teams have a moderate level of cultural diversity; in addition, successful teams engage in collaborative activities with teams from other European countries and the US leading to joint publications" (p. 23).

Besides, there are studies which also consider analyzing the scientific collaboration among researchers at institute and department levels (Abbasi, Hossain, Uddin, & Rasmussen, 2011). The precise role of geographical proximity for research

Table 1
Top 10 cited journals in “Information Science & Library Science”.

Rank	Journal name	2009 impact factor
1	MIS Quarterly	4.49
2	Journal of American Medical Informatics Association	3.97
3	Journal of Computer-Mediated Communication	3.64
4	Journal of Informetrics	3.38
5	Annual Review of Information Science & Technology	2.93
6	International Journal of Computer-Supported Collaborative Learning	2.69
7	Journal of American Society for Information Science & Technology	2.30
8	Information & Management	2.28
9	Journal of the Association for Information Systems	2.25
10	Scientometrics	2.17

collaboration and knowledge exchange still remains unclear (Ponds et al., 2007). For example, Katz (1994) stated lack of evidence on whether or not geographical proximity had a significant influence on scientific cooperation.

3. Data and methods

Using a text mining application on the one-to-one co-author pair affiliation field ‘authors with affiliations’ (provided by Scopus), distinct number of affiliations (departments), institutes and countries are retrieved for each publication. Therefore, each publication’s frequency of intra- and inter-departmental, intuitional and national collaborations among their authors is identified. In addition, normalizing publications citation counts (average number of citations a publication receives per year) is used as a surrogate for publication impact or quality. Finally, a correlation test is applied to examine to what extent each of these different levels of internal and external collaborations associate with the publication’s impact.

3.1. Sample data and methodology

For the analysis, publication data of top 10 journals of the “Information Science & Library Science” were selected based on their 2009 impact factor as enumerated in the current edition of *Journal Citations Report*.[i] Publication meta-data of these selected journals between 2000 and 2009 inclusive is extracted, as listed in Table 1, from Scopus (an online bibliometric database) by considering just articles and review articles for which their author(s) and affiliation records were available. While the list includes top 10 journals but unfortunately, we could not get the data of the “*Journal of the Association for Information Systems*” (Rank 9).

The following simplification or rationalization actions were taken prior to analysis of the data: (1) publications meta-data extraction; (2) data verification and cleaning; (3) processing to data to record collaboration measures.

After extracting the publications meta-data, publication information (i.e., title, publication date, author names, affiliations, publisher and number of citations) is stored in a relational local database. Different types of information were extracted from each publication meta-data: publications information (i.e., title, publication date and journal name); authors’ names and affiliations (including country, institution and department names).

Since this study conducts different macro-, meso- and micro-levels of analysis (i.e., country, institute, department), affiliation data is an important consideration in this research. But the affiliation information is found to be messy in the original extracted data (e.g., different written name for a country, institute and department) and even for some of the publications the data for ‘country’ field was blank or mixed with ‘institute’ field. So, in the second phase a manual checking (using Google) is done to fill the missing fields using other existing fields (e.g., institute names is used to find country). Finally, after cleaning of the data, using a text mining application on the affiliation field, the distinct number of affiliations (departments), institutes and countries are identified for each publication. The resulting database contained 4837 publications that received totally 93,660 citations, reflecting the contributions of 8069 distinct authors.

To quantify collaboration activities, ‘authors with affiliations’ field, a one-to-one affiliation–author field in Scopus dataset, was processed and the distinct number of collaborations for each publication was extracted. In order to count the number of affiliations for each level, the extracted number of distinct affiliations, the number of distinct institutes (e.g., universities) and the number of countries for each publication were used. To determine the internal and external collaboration count, the number of pairs of authors (co-authors) who have internal (intra-departmental: both authors from the same department of an institute) or external (inter-departmental: both authors from different departments, no matter from the same institute and country or not) collaboration was counted. Also, the internal and external collaborations counts were determined for the institute and the country level similar to the department level.

As an illustration, if a publication has ‘*A*’ number of authors, ‘*D*’ number of distinct departments, ‘*I*’ number of distinct institutes, and ‘*C*’ number of distinct countries, its total number of collaborations is $\{A \times (A-1)/2\}$. The publication’s intra-departmental, inter-departmental, intra-institutions, inter-institutions, intra-national, and international collaboration

[i] http://sciencewatch.com/dr/sci/10/aug15-10_1.

Table 2
Sample publications' multi-level collaboration count.

Publication ID	P1	P2	P3	P4	P5
Number of authors	3	9	2	2	1
Number of departments	3	5	1	2	1
Number of institutes	2	5	1	2	1
Number of countries	2	3	1	1	1
Number of collaborations	3	36	1	1	0
Number of intra-departmental (departments internal) collaborations	0	10	1	0	0
Number of inter-departmental (departments external) collaborations	3	26	0	1	0
Number of intra-institutional (institutions internal) collaborations	1	10	1	0	0
Number of inter-institutional (institutions external) collaborations	2	26	0	1	0
Number of intra-national (country internal) collaborations	1	21	1	1	0
Number of inter-national (country external) collaborations	2	15	0	0	0

count are calculated as $\{(A-(D-1))\times(((A-(D-1))-1))/2\}$, if $(A>D)$ $\{((A-(D-1))\times(D-1))+((D-1)\times(D-2)/2)\}$ otherwise $\{D\times(D-1)/2)\}$, $\{(A-(I-1))\times(((A-(I-1))-1))/2\}$, if $(A>\mathrm{I})$ $\{((A-(I-1))\times(I-1))+((I-1)\times(I-2)/2)\}$ otherwise $\{I\times(I-1)/2)\}$, $\{(A-(C-1))\times(((A-(C-1))-1))/2\}$, and (intra-departmental–intra-national) respectively.

Table 2 shows a sample dataset; for instance, P1 has 3 authors from 3 different departments which are located in 2 different institutes, each institute being in a different country. Therefore, the number of collaborations for P1 is 3. There are no intra-departmental collaborations but there are 3 inter-departmental collaborations as 3 authors are from 3 different departments. Since 3 authors are from 2 institutes, it means that 2 of them are from similar institutes meaning that there is 1 intra-institutions collaboration and 2 inter-institutions collaborations. A similar logic is applied for the country level.

For all publications, the frequency of intra- and inter-departmental, intuitional and national collaborations among authors are identified. In order to make the impact (quality) of publications published in different years comparable, their citation count is normalized. A normalized citation count of a publication is equal to the number of citation counts received divided by its age. The age is found by subtracting the publication year from 2011, being the year the citation count data were extracted.

3.2. Measures

3.2.1. Research impact indicator

The number of citations a publication receives reflects how it has influenced or has attracted the attention of other researchers. Obviously publications that have published earlier had more chance to get more citations. Therefore, in order to be able to compare impact of publications which have been published in different years, the number of citations are normalized considering how long has been passed from the year it has been published. Thus, to evaluate the impact of a publication an average number of citations a publication receives per year (as a normalization measure) is computed.

3.2.2. Geographical (physical proximity) collaboration indicators

In order to measure the research collaborations activity, the number of collaborations between authors of a publication is used. To measure internal collaboration, the number of collaborations among the authors within the same unit for each level of analysis (department, institution and country) is counted. And to consider external collaborations, the number of collaborations among authors from different departments, institutions and countries for each level of analysis is counted. For instance, each pair of co-authors from the same departments is counted as the number of intra-departmental collaborations (departments' internal collaborations) and the frequency of pair of co-authors from different departments is counted as the number of inter-departmental collaborations (departments' external collaborations) and the like for the institutions and countries of origin.

3.3. Research model

Fig. 1 demonstrates the research model which is used to investigate the association between publication impact and authors' collaboration activity. To compare the different impact of the collaboration among the authors inside the same unit (internal collaboration) and the authors out of the unit (external collaboration) by considering network data, internal and external co-authors are counted.

Based on the research model, the following hypotheses are proposed:

(H1): Publication's total collaboration count correlates to its impact.
(H2a): Publication's intra-departmental collaboration count correlates to its impact.
(H2b): Publication's inter-departmental collaboration count correlates to its impact
(H3a): Publication's intra-institutions' collaboration count correlates to its impact.
(H3b): Publication's inter-institutional collaboration count correlates to its impact
(H4a): Publication's intra-national collaboration count correlates to its impact.
(H4b): Publication's inter-national collaboration count correlates to its impact.

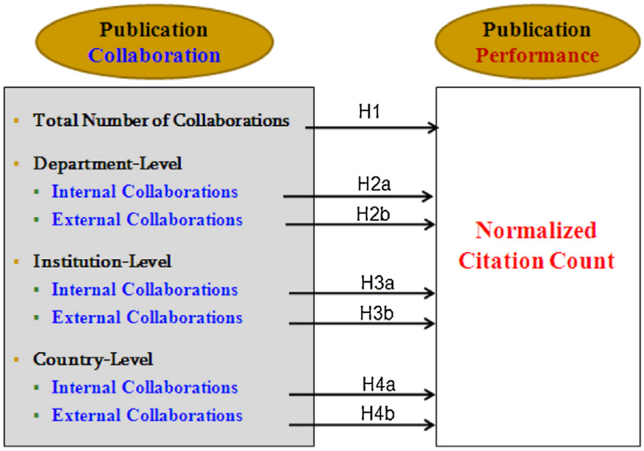

Fig. 1. Research model.

Table 3
Top journals publications productivity and impact.

	Journal name	# of pub.	Sum of cit.	Sum of norm. cit.	Avg. norm. cit./pub.	2009 imp. factor rank
1	MIS Quarterly	279	20,410	2958.09	10.60	1
2	Information and Management	585	16,444	2512.79	4.30	8
3	J. of the American Medical Inf. Association	768	17,306	2977.63	3.88	2
4	J. of Informetrics	99	982	320.05	3.23	4
5	Annual Review of Inf. Sci. & Technology	96	1597	278.04	2.90	5
6	J. of Computer-Mediated Communication	406	5636	1080.33	2.66	3
7	Int. J. of Computer-Supported Collab. Learning	89	766	211.38	2.38	6
8	J. of the American Society for Inf. Sci. & Tech.	1383	18,389	3153.52	2.28	7
9	Scientometrics	1132	12,130	2177.96	1.92	10

4. Results

4.1. Journals' publications productivity and impact statistics

Table 3 lists the journals which are selected for this analysis and their number of publications (between 2000 and 2009), the sum of citations they received, the sum of publications' normalized citation counts, and the average of normalized citations per publication followed by their rank based on 2009 impact factor as the last column.

Table 3 indicates that the subject journals did not publish the same amount of publications; and their citation counts vary. In order to compare these journals properly, the average number of citations per year per publication is calculated for each one as shown in the fifth column. The subject journals are sorted in Table 3 based on this column and as seen, "*MIS Quarterly*" publications received more citations on average (10.60 average citations per year per publication) followed by "*Information and management*" (4.30) with a big gap. Obviously, the new ranks (which are based on the citations received by 2011 for the articles published between 2000 and 2009) should not be the same as the 2009 Impact Factor (which is based on the citations received at 2009 to the journal's articles published during 2007–2008) but as displayed in Table 3, there is not a big difference except for "*Information and Management*". This may be owing to more citations "*Information and Management*" publications received after 2008.

4.2. Publications type's impact statistics

Only two main types of publications are considered: articles and review articles. Table 4 presents the number of publications, the sum of citations they received, the average citations per publication, the sum of publications' normalized citation counts and the average of normalized citations per publication for each publication type. As the result reveals, while the number of articles are much more than the review articles but the average citations received per publication and also average

Table 4
Publications type impacts.

Publication type	# of publications	Sum of citations	Avg. cit./pub	Sum of norm. cit.	Avg. norm. cit./pub.
Article	4266	78,040	18.29	13,356.23	3.13
Review	571	15,620	27.36	2313.56	4.05

Table 5
Publications counts and their citations statistics over 10 years.

Year	# of publications	Sum of citations	Avg. cit./pub	Sum of norm. cit.	Avg. norm. cit./pub.
2000	325	9936	30.57	901.9	2.78
2001	332	11,246	33.87	1124.6	3.39
2002	352	10,278	29.20	1140.5	3.24
2003	381	12,875	33.79	1608.4	4.22
2004	397	11,653	29.35	1663.0	4.19
2005	484	10,109	20.89	1683.2	3.48
2006	589	10,406	17.67	2081.2	3.53
2007	643	9091	14.14	2272.8	3.53
2008	607	5021	8.27	1671.7	2.75
2009	727	3045	4.19	1522.5	2.09
Total	4837	93,660	19.36	15,670	3.24

citation counts per year per publication of review articles are higher than other articles. Our results indicate that a review article receives on average 4.05 citations per year while a non review article just receives 3.13 citations per year.

4.3. Publications impact statistics over time

Table 5 indicates the number of publications, the sum of citations received, and the sum of the normalized citations count of all publications per year. As listed, more publications were published each year compared to the previous year and this number becomes more than double from 325 publications in 2000 to 727 publications in 2009. The yearly sum of citations (received by 2011 that the data were extracted) does not show a linear growth. Although older publications have higher chance of receiving citations compared to the new ones but the increasing number of publications over time also gets more chance of receiving higher citations to recent years. Therefore, it is not a suitable indicator to compare the impact of yearly publications. The average number of citations per publication (shown in fourth column) reveals a decreasing trend over time but again it is not a linear line and publications published in 2001 received more citations on average followed by the ones published in 2003 and 2000.

Considering the publications' sum of normalized citation counts (average sum of citations per year until 2011 which is the sum of citations divided by (2011 – 'Year')) in Table 5, an increasing linear trend up to 2007 is seen followed by a linear decline for 2008 and 2009. As the number of publications per year varies, to properly compare each year's publication impact, the publications' sum of normalized citation counts is divided by the number of publications per year. The results (column 6) indicate a very fluctuating trend by having the highest average citation counts per year for 2003 publications followed by 2004, 2006 and 2007 publication (4.22, 4.19, 3.53 and 3.53 citations respectively) and the lowest value for 2009, 2008 and 2000 publications (2.09, 2.75 and 2.78 citations respectively). Therefore, the result demonstrates that not necessarily older publications will receive more citations and there is a peak period for attracting more citations. To find the precise threshold values, access to annual citations count and extra analysis are required, which is not the focus of this study. Based on the citations data of 2011 for the selected publications, it could be inferred that in this domain articles attract more citations between 7 and 8 years after publication.

4.4. Publications' collaborations statistics

Table 6 shows more information about the distribution of the number of authors per publication and also the collaboration counts based on different affiliation levels. It also demonstrates the average normalized citation counts per publication. Having 2.64 authors per publication (on average) from 2.11 departments and 1.84 institutes and 1.33 countries reflects a reasonably good picture of external collaborations among researchers in "Information Science & Library Science".

First part of Table 6 lists the majority of publications in the dataset with two authors (32%) followed by single-author publications (27%) and then three-author publications (20%). As shown the more authors' publications have, more citations on average they receive. Using average normalized citation counts per publication also indicates that a 4-author publication receives on average 3.98 citations per year followed by 3-author, more than 4-author and 2-author publications and finally single-author publications receive the least rank by only 2.45 citations per year on average.

Although the second part indicates 37% of publications are result of intra-departmental collaborations but still a good portion of publications' are the result of inter-departmental collaborations. Again the result reveals more citation counts per year for a publication that has more authors from different departments. Also, the result shows that about half of the publications' authors are in the same institute and 74% of them are due to an intra-national collaboration. But publications' impact is higher for publications that have more geographically diverse authors from different institutes and countries. Thus, high positive correlation between publication impact and its number of collaborations especially external collaborations can be anticipated.

Table 6
Publications' collaborations distributions.

	# of pub.	Sum of norm. citation	Avg. nor. cit. per pub
Avg. number of authors/paper: 2.64			
1-Author papers	27%	20.5%	2.45
2-Author papers	32%	33.4%	3.38
3-Author papers	20%	21.1%	3.48
4-Author papers	10%	12.2%	3.98
>4-Author papers	11%	12.8%	3.63
Avg. number of departments/paper: 2.11			
1-Department papers	37%	31.2%	2.73
2-Department papers	35%	33.5%	3.11
3-Department papers	16%	19.1%	3.76
4-Department papers	8%	10.7%	4.60
>4-Department papers	4%	5.5%	4.40
Avg. number of institutes/paper: 1.84			
1-Institute papers	48%	38.1%	2.60
2-Institute papers	32%	34.0%	3.41
3-Institute papers	13%	16.7%	4.16
>3-Institute papers	7%	11.2%	5.05
Avg. number of countries/paper: 1.33			
1-Country papers	74%	68.3%	3.00
2-Country papers	21%	23.9%	3.64
>2-Country papers	5%	7.9%	5.02

Fig. 2 depicts in detail the distribution of the proportion of publications and the number of authors (a), the number of departments (b), the number of institutes (c) and the number of countries (d). As Fig. 2 displays, the majority of the publications have few authors and very few publications have multiple authors. This indicates that the "*Information Science and Library Science*" collaboration networks most probably follow a power-law probability distribution.

4.5. Correlation

In order to examine the proposed hypotheses, Pearson correlation test is run and as demonstrated, in Table 7, all the geographical (physical proximity) collaboration measures (i.e., the number of collaborations both internally and externally in

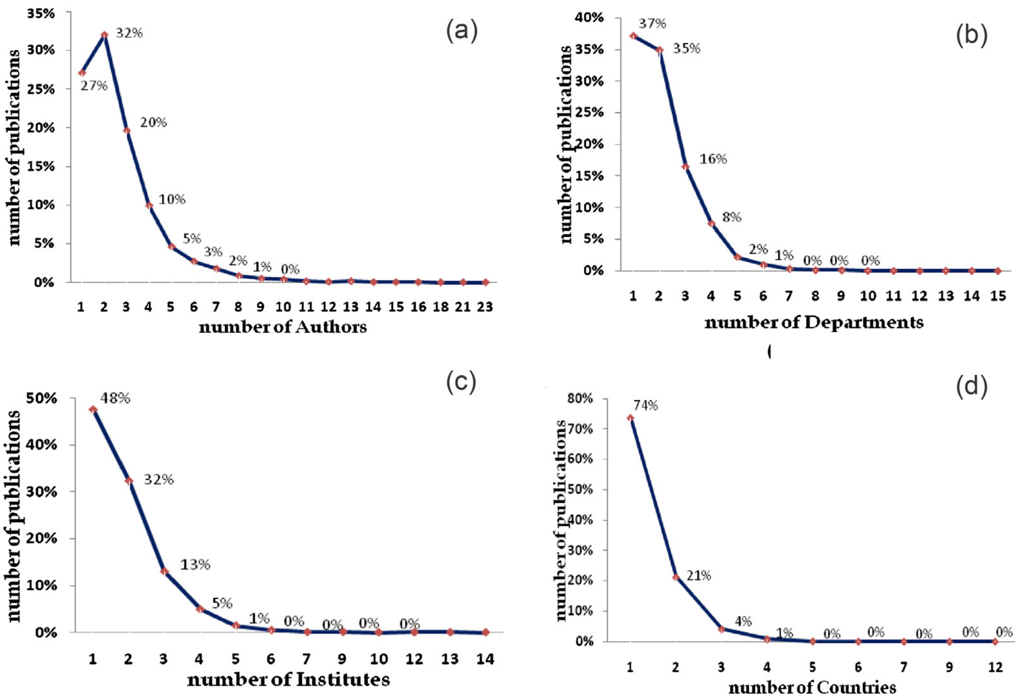

Fig. 2. Distributions of number of publications and their authors' different level of affiliations.

Table 7
Collaboration measures' correlations coefficient.

Correlations coefficient (N = 4837) Publication collaborations measures		Publication impact Normalized citations count
Total collaboration count	*Rho*	.054[*]
Intra-departmental collaboration count	*Rho*	.033[*]
Inter-departmental collaboration count	*Rho*	.061[*]
Intra-institutional collaboration count	*Rho*	.029[*]
Inter-institutional collaboration count	*Rho*	.067[*]
Intra-national collaboration count	*Rho*	.048[*]
Inter-national collaboration count	*Rho*	.042[*]

[*] Correlation is significant at the .01 level (2-tailed).

all levels) show a positive and significant correlation with the publication impact measure (the average citations per year). Total collaboration counts are highly correlated to the publication impact. Comparing the internal and external collaborations association with the publication impact, the result indicates that collaborations cross departments and institutions (inter-departments and inter-institution collaborations) have much higher correlation coefficients compared to the domestic (internal) collaborations to the publication impact but considering the country level collaborations, internal collaborations have higher associations rather than international collaborations. Thus, it can be inferred that the publications with more authors from different departments and different institutions but inside the same country will have better impact and receive more citations on average.

5. Conclusion and discussion

This study provides insights into the relationship between scientific collaboration and research impact. The approach adopted aimed at addressing this topic by considering both internal and external collaborations in different affiliation levels (i.e., departments, institutes and counties), by including papers from a wide variety of countries. Furthermore, the results of this macro-, meso- and micro-level analysis could be taken into account when deciding with whom to collaborate through using the combination of internal or external collaborators for each affiliation level.

In order to address whether the collaboration counts and the research impact (or quality) are associated, the data related to research articles and review articles published in the top 10 high impact journals of "*Information Science & Library Science*" were retrieved from Scopus. Data sets were classified based on different types of geographical (physical proximity) collaboration metrics of authors (i.e., internal and external collaborations based on their affiliations). The normalized citation counts (i.e., average number of citations per year) were used as a surrogate for evaluating the publication impact.

The results reveal that research impact is positively associated with all levels of collaboration metrics. Thus, in general, the publications with more diverse authors (collaborations) have better impact especially if the collaborations are more external rather than internal, considering their affiliations' departments and institutes but more internal collaborations considering their country. The fact that international collaboration has a lower correlation to publications' impact may be due to the apparent challenge of collaboration across national and cultural boundaries. The reason for intra-departmental collaboration's low correlation to publications' impact may be explained by exchanging redundant knowledge among the researchers in the same departments (as usually have access to similar kinds of resources and equipment).

Therefore, the findings support that having co-authors with diverse knowledge and skills enhance scholars' knowledge and experience through decreasing the research project process, including writing and revision process of publication (as the output of the work) and also improving the impact.

It is also found that while co-authored publications have better impact, there is a threshold for the number of authors and in this dataset it is found that publications with four authors have been cited more than other publications. Having access to more personal information of researchers (i.e., age and sex) could be useful moderating variables in the proposed model. In that case, different attributes could be categorized to examine the different effect for each category, which could generalize the model. Lack of access to this kind of information can be considered as one of the limitations of this research.

It is also known that the extent to which researchers co-author varies among scientific fields and it is usually assumed that this is caused by variation in the level of collaboration. To investigate how scientific fields (due to their different collaboration characteristics) influence on the association of researchers' collaborations activities and their performance, it is necessary to conduct similar analysis for several research collaboration groups from different fields as a future work.

References

Abbasi, A. (2013). h-Type hybrid centrality measures for weighted networks. *Scientometrics*, 1–8.

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.

Abbasi, A., Chung, K. S. K., & Hossain, L. (2012). Egocentric analysis of co-authorship network structure, position and performance. *Information Processing & Management*, *48*(4), 671–679.

Abbasi, A., Hossain, L., & Owen, C. (2012). *Exploring the relationship between research impact and collaborations for Information Science*. Paper presented at the 2012 45th Hawaii international conference on system science (HICSS).

Abbasi, A., Hossain, L., Uddin, S., & Rasmussen, K. J. R. (2011). Evolutionary dynamics of scientific collaboration networks: Multi-levels and cross-time analysis. *Scientometrics*, *89*(2), 687–710.

Ahuja, M. K., Carley, K., & Galletta, D. F. (1997). Individual performance in distributed design groups: An empirical study. In *Proceedings of the 1997 ACM SIGCPR conference on Computer personnel research* ACM, (pp. 160–170).

Avkiran, N. (1997). Scientific collaboration in finance does not lead to better quality research. *Scientometrics*, *39*(2), 173–184.

Barjak, F., & Robinson, S. (2008). International collaboration, mobility and team diversity in the life sciences: Impact on research performance. *Social Geography*, *3*(1), 23–38.

Borgman, C., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, *36*(1), 2–72.

Cross, R., & Cummings, J. (2004). Tie and network correlates of individual performance in knowledge-intensive work. *The Academy of Management Journal*, *47*(6), 928–937.

Cummings, J., & Cross, R. (2003). Structural properties of work groups and their consequences for performance. *Social Networks*, *25*(3), 197–210.

Cummings, J. N. (2004). Work groups, structural diversity, and knowledge sharing in a global organization. *Management Science*, *50*(3), 352–364.

Demsetz, H. (1991). The theory of the firm revisited. In O. E. Williamson, S. G. Winter, & R. H. Coase (Eds.), *The nature of the firm: Origins, evolution, and development* (pp. 159–179). New York: Oxford University Press.

Duque, R., Ynalvez, M., Sooryamoorthy, R., Mbatia, P., Dzorgbo, D., & Shrum, W. (2005). Collaboration paradox: Scientific productivity, the Internet, and problems of research in developing areas. *Social Studies of Science*, *35*(5), 755.

Eisenhardt, K., & Tabrizi, B. N. (1995). Accelerating adaptive processes: Product innovation in the global computer industry. *Administrative Science Quarterly*, *40*(1)

Georghiou, L. (1998). Global cooperation in research. *Research Policy*, *27*(6), 611–626.

Gerrity, D. M., & McKenzie, R. B. (1978). The ranking of southern economics departments: New criterion and further evidence. *Southern Economic Journal*, 608–614.

Glänzel, W., Schubert, A., & Czerwon, H.-J. (1999). A bibliometric analysis of international scientific cooperation of the European Union (1985–1995). *Scientometrics*, *45*(2), 185–202.

Gomez, I., Teresa Fernández, M., & Sebastian, J. (1999). Analysis of the structure of international scientific cooperation networks through bibliometric indicators. *Scientometrics*, *44*(3), 441–457.

Griffin, A., & Hauser, J. R. (1992). Patterns of communication among marketing, engineering and manufacturing – A comparison between two new product teams. *Management Science*, *38*(3), 360–373.

Hoekman, J., Frenken, K., & Van Oort, F. (2009). The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, *43*(3), 721–738.

Jiang, Y. (2008). Locating active actors in the scientific collaboration communities based on interaction topology analyses. *Scientometrics*, *74*(3), 471–482.

Katz, J., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, *40*(3), 541–554.

Katz, J., & Martin, B. (1997). What is research collaboration? *Research Policy*, *26*(1), 1–18.

Katz, J. S. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, *31*(1), 31–43.

Laband, D. N. (1985). An evaluation of 50 "Ranked" economics departments: By quantity and quality of faculty publications and graduate student placement and research success. *Southern Economic Journal*, 216–240.

Laudel, G. (2002). What do we measure by co-authorships? *Research Evaluation*, *11*(1), 3–15.

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, *35*(5), 673.

Lehmann, S., Jackson, A., & Lautrup, B. (2006). Measures for measures. *Nature*, *444*(7122), 1003–1004.

McFadyen, M., Semadeni, M., & Cannella, A. A. (2009). Value of strong ties to disconnected others: Examining knowledge creation in biomedicine. *Organization Science*, *20*(3), 552–564.

Melin, G. (2000). Pragmatism and self-organization: Research collaboration on the individual level. *Research Policy*, *29*(1), 31–40.

Melin, G., & Persson, O. (1996). Studying research collaboration using co-authorships. *Scientometrics*, *36*(3), 363–377.

Monge, P., Rothman, L., Eisenberg, E., Miller, K., & Kirste, K. (1985). The dynamics of organizational proximity. *Management Science*, *31*(9), 1129–1141.

Nankani, E., Simoff, S., Denize, S., & Young, L. (2009). Supporting strategic decision making in an enterprise university through detecting patterns of academic collaboration. *Information Systems: Modeling, Development, and Integration*, 496–507.

Narin, F. (1991). Globalization of research, scholarly information, and patents – Ten year trends. *The Serials Librarian*, *21*(2), 33–44.

Pinto, M. B., Pinto, J. K., & Prescott, J. E. (1993). Antecedents and consequences of project team cross-functional cooperation. *Management Science*, *39*(10), 1281–1297.

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, *86*(3), 423–443.

Sonnenwald, D. (2007). Scientific collaboration: A synthesis of challenges and strategies. *Annual Review of Information Science and Technology*, *41*, 643–681.

Stigler, G. J., & Friedland, C. (1975). The citation practices of doctorates in economics. *The Journal of Political Economy*, *83*(3), 477–507.

Stokols, D., Harvey, R., Gress, J., Fuqua, J., & Phillips, K. (2005). In vivo studies of transdisciplinary scientific collaboration: Lessons learned and implications for active living research. *American Journal of Preventive Medicine*, *28*(2), 202–213.

van Raan, A. (2004). Measuring science. Capita select of current main issues. In *Handbook of quantitative science and technology research. The use of publication and patent statistics in studies of S&T systems*.

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.

Yan, E., Ding, Y., & Zhu, Q. (2010). Mapping library and information science in China: A coauthorship network analysis. *Scientometrics*, *83*(1), 115–131.