An Analysis Model to Characterize the Talent Group Composition of Research Institute

Jing Xu[1, 2], Lina Wang[1]
[1]Chengdu Library and Information Center, Chinese Academy of Sciences
Chengdu, 610041
[2]University of Chinese Academy of Sciences
Beijing, 100049

ABSTRACT: *Carrying out study on talent group composition can help us insight its present situation, and discover its potential problem, such as population aging, knowledge flow barrier, inbreeding, etc. An analysis model is proposed to characterize the talent group composition of research institute in three dimensions: age distribution, education experiences and collaboration pattern. Moreover, the analysis is based on objective data, and objective indicators. Age distribution may reflect the population aging problem whether or not; education experiences may reflect the inbreeding problems; while collaboration pattern may reflect the knowledge flow problems. According to the indicator of each dimension, research institutes are divided into different types, the classification reveals different problems they may have. At last, an empirical study was conducted on two research institutes to prove the effectiveness of this model.*

Keywords: Research Institute, Talent Group Composition, Age Distribution, Education Experiences, Collaboration Pattern

Received: 4 December 2019, Revised 8 March 2020, Accepted 8 April 2020

DOI: 10.6025/jstm/2020/1/2/71-76

1. Introduction

As an important social activity, scientific research is the foundation for scientific promoting and technological progress for social and economic development. The most critical element is talent group. A lot of researches have been conducted on exploring knowledge sharing, innovation performance and collaboration pattern, and so on. The results show that the talent group composition has a great impact on research activities [1-5]. Early in 1980s, Chinese professor Liu[6] has pointed out the importance of optimizing the composition of scientific research talent group, that without suitable disciplinary distribution, skill level distribution and age distribution, the overall effectiveness and performance of the talent group could be reduced.

More and more organizations have begun to pay attention to the composition of talent group and how to optimize the composition for their own organization. Countries and regions also emphasize the importance of talent group optimization. Thus it can be seen that the research on talent group composition is an urgent problem to be solved. The talent group problem is a complex system, with multiple levels and multiple factors, including discipline distribution, intelligence level distribution, age distribution, etc. The purpose of solving this problem is to promote the talent flow, avoiding academic inbreeding, remaining innovation energy, promoting knowledge renewal, etc. Carrying out study on talent group composition of scientific research institute can help us insight its present situation, and discover its potential problem, such as population aging problem, knowledge flow barrier, academic inbreeding, etc. Institutes can find their problems and make adjustments on talent group timely, according to the analysis results. Therefore, it is of great theoretical and practical significance to study on this problem.

Up to now, there are few discussions on the composition of talent group. Related research mainly focuses on the scientific team, including team demographic characteristics, team heterogeneity, the relationships between team structure and team performance, as well as inter-group relationships within scientific team. Besides, the demographic characteristics of the team mainly refer to age, gender, race or nation, cultural background, etc. For the heterogeneity of the team, the scope is broader, and there is no uniform definition. Many researchers explored the positive and negative impacts of team demographics, team heterogeneity, and team size on the scientific team from different perspectives. A prior research on 89 teams from 1,415 experiments and 2,128 publications has observed the following patterns: disciplinary diversity has a positive effect on team productivity; gender diversity has no obvious connection with team productivity; and seniority diversity has a negative interaction on team productivity [7]. Some research has shown that distance also influences team performance. Pelled[8] found that there is a significant positive correlation between the diversity of subject background and the team performance. Paul[9] found that teams with diverse disciplines are more likely to inspire innovation than teams with single-disciplinary background. Of course, some researchers believe that the differences between members will lead to communication barriers, which are harmful to collaboration [10].

Research on talent group composition of institute is rare up to now, thus there is no method to discover possible problems in talent group. This paper lists the possible problems that may exist in institutes, organizations, such as ageing problem, academic inbreeding, knowledge flow problems, and then proposes a three dimension analysis mode based on age distribution, education experiences, and collaboration pattern to evaluate the status of institutions.

The following paper is structured as follows: chapter 2 introduces a three-dimensional analysis model, including the research purpose, method, indicators and feature types of each dimension in detail. Chapter 3 presents the empirical analysis results on two research centers in the field of quantum information science respectively from China and the United Kingdom. Finally the conclusion section summarizes the main contributions of this study, reviews the limitations and gives some suggestions for future research on this topic.

2. Definition and Methodology

2.1. Definition of Talent Group Composition

The object in this study is limited to scientific research institute, such as research centers, laboratories, etc., and the element of the object is every single person. Considering that people with mobility (such as graduate students, doctoral students, postdocs, etc.) less influence on institute than Long-term employee, this paper only considers employees, who are engaged in scientific research, such as senior researchers, associate researchers, professors, and associate professors.

2.2. Age Distribution

Purpose: institute with suitable age distribution can stimulate innovation and promote internal mobility in institute, while avoiding population aging problems. Studies have demonstrated that the scientist has golden age period, about 35-50 years old [11]. Some research has found out that the average age of Nobel Prize winners at the time that completed their studies is 38.7 years old [12].

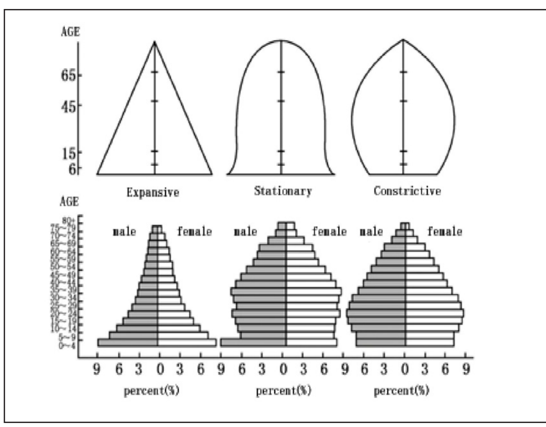

Figure 1a. Three types of age pyramid in demographics

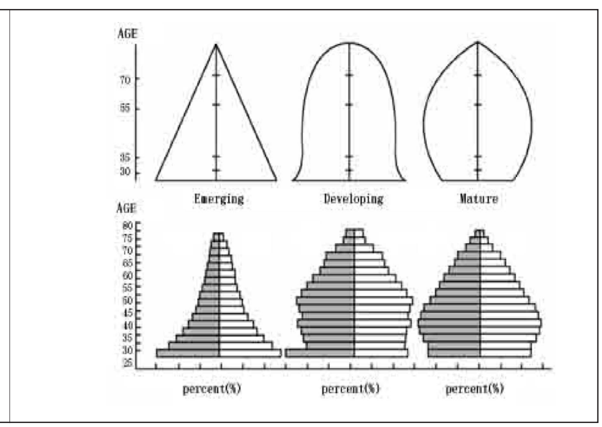

Figure 1b. Three types of age Pyramid in talent group composition problem

Indicator: Age Pyramid

Types: Emerging; Developing; Mature.

Method: Age Pyramid in demographics is introduced into this model to classify talent group into three types: ‘Emerging: People in this institute are relatively young, mainly under 40 years old, with great potential in academic research;

Developing: People in this institute are mainly middle-aged, most of whom are in their 40s and 50s, with strong scientific research ability and promising development trends; ƒMature: This institute is relatively stable, that most members are over 50 years old. The institute may face the problem of limited space for development. (As in Figure 1).

2.3. Education Experiences

Purpose: We can find out the background similarity of members in talent group, and evaluate the possibility of academic distant hybridization by discussing education experiences. In fact, the background of education experience of each person is one of the heterogeneity of talent group, which can reflect the academic consanguinity structure, such as teacher-student, alumnus. From a sociological perspective, a diverse background can lead to more diverse external collaborations and communications.

Indicator: Network density of education experiences network.

Types: Heterogeneous; homogenous.

Method: Each node represents a member, and if two members used to study in the same organization (the education experiences of the members refers to the academic degrees. As for the Ph.D. researchers’ education experiences in this study refer to the schools that he/she get his/her undergraduate degree, master degree, or doctoral degree.), there will be an edge between the two nodes, thereby a network can be built. And the network density is used as the measurement indicator.

For an undirected and unauthorized network $G(E, V)$, representing education experiences network. E is the node set, representing all members of talent group in one institute; V is the edge set, representing all relationships between these members. The network density of G graph is $d(G)$, where N is the number of network nodes, and L is the number of links in graph G.

$$d(G) = \frac{2L}{N(N-1)}$$

The network density reflects the similarity between members. According to this indicator, the talent group can be classified into two types: heterogeneous and homogeneous. According to Facebook data base (data base from http://snap.stanford.edu/data/ego-Facebook.html), the density of Facebook network is 0.1, which means the network density of normal social network is around 0.1. While some research shows that the maximum density proved in actual network is 0.5. The education experiences network defined in this study is more similar to the social network. Two members in one institute may have the same education experiences but no contact, so the network density is higher than that of the normal social network, the demarcation value of the network density is set to 0.3. If the network density is greater than or equal to 0.3, and the larger the value, the higher the similarity of education experiences, the more homogenous the talent group tends to be, and this indicates an academic inbreeding phenomenon; Conversely, if the network density is less than 0.3, the smaller the value, the lower the similarity of the education experience, the more heterogeneous the talent group tends to be.

2.4. Collaboration Pattern

Purpose: Collaboration pattern can reflect the depth and breadth of collaboration among members, and evaluate the knowledge flow in this institute. Relevant studies have explored the relationship between collaboration network and innovation. It has been shown that institutes with high research outputs have closer research collaboration networks, but excessive high cohesion may hinder new opportunities for innovation [13]. By analyzing the role of weak correlation in knowledge transfer between sub-units of an institution, Hansen [14] found that weak relationships can help members explore useful information of other sub-units, but hinder the transfer of complex knowledge.

Indicator: Average path length.

Types: Cohesive; Flexible.

Method: Based on social network theory, this study builds a research collaboration network based on co-authorship of papers. For an undirected and unauthorized network $G(E, V)$, representing collaboration network. E is the node set, representing all members of talent group in one institute; V is the edge set, representing all co-authorships between these members. The average path length of G graph is AL, the average distance between all two nodes in the network.

$$AL = \frac{1}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} d_{ij}$$

N is the number of network nodes, d_{ij} indicates the distance between node i and node j. And we take the average path length as an indicator to measure the degree of scientific collaboration. It can describe the length of the communication links between network nodes as a whole, as it is an important indicator to describe the performance and efficiency of network transmission[15]. In the perspective of knowledge diffusion, the smaller the AL, the shorter the average distance experienced by knowledge passing from one node to another in the network, indicating that the network transmission efficiency is higher[16]. The AL of the collaboration network reflects the knowledge flow in the talent group. Based on this indicator, the talent group can be classified into cohesive and flexible. Bactstrom and etc.[17] did some research on Facebook network, with 7,210 million users and 69 billion friendship connections, the average path length they calculated is 4.74. If the average path length is greater than or equal to 4.74, the larger the value, the more cohesive the talent group, which is conducive to the communication among members. On the contrary, if the average path length is less than 4.74, the smaller the value, the more flexible the collaboration and the lower information redundancy.

2.5. Three-dimensional Analysis Model

Considering the characteristics of the talent group, the objectivity and accessibility of the research data, this study proposes a three-dimensional analysis model in Table 1.

Dimension	Indicator	Types	Potential Problems
Age Distribution	Age Pyramid	Emerging; Developing; Mature	Population aging problem
Education Experiences	Network Density in Education Network	Heterogeneous; Homogenous	Academic inbreeding
Collaboration Pattern	Average Path Length in Collaboration Network	Cohesive; Flexible	Knowledge flow problems

Table 1. Three-dimensional analysis model

3. Empirical Study

We choose two research centers on quantum science from China and England, which is Centre for Excellence in Quantum Information and Quantum Physics of the Chinese Academy of Sciences (CAS-QC) and Centre for Quantum Computation of Oxford University (Oxford-Q).

Founded in 1998, Oxford-Q is one of the largest centers for quantum science research in the world, with 38 independent research groups and 43 full-time researchers. While CAS-QC was established in 2014, gathered a great number of high level talents. There are 62 researchers engaged in scientific research.

The data comes from their official website, personal homepage and the Web of Science database. The professor list of CAS-QC comes from email consultation. The paper datasets are SCI papers before 2018 (retrieval time: January 2019).

3.1. Age Distribution

The members' age in CAS-QC are mainly concentrated in 30-39 years old, it can be categorized as the emerging, that have great development potential in the future. For Oxford-Q, whose most members are in their golden age of development can be regarded as developing (see Table 2 and Figure 2).

Institutes	Age ratio					Types
	Age 30-39	Age 40-49	Age 50-59	Age 60-69	Age 70-79	
CAS-QC	47.62%	31.75%	19.05%	0.00%	1.59%	Emerging
Oxford-Q	24.44%	42.22%	22.22%	13.33%	0.00%	Developing

Table 2. Age distribution of CAS-QC and Oxford-Q

3.2. Education Experiences

The results show that the network density of CAS-QC and Oxford-Q is 0.43 and 0.54 respectively, and indicate they both belong to homogeneous, which means academic inbreeding phenomenon exists in both centers (see table 3).

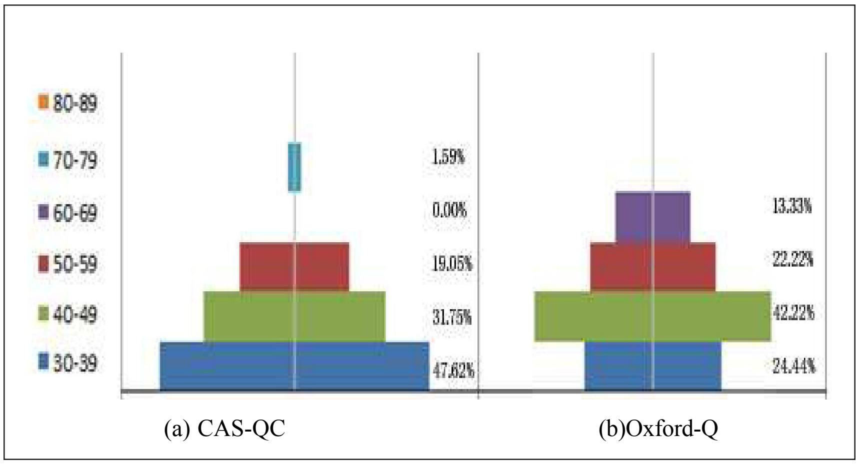

Figure 2. Pyramid age of CAS-QC and Oxford-Q

Institutes	Nodes	Links	Network Density	Types
CAS-QC	62	813	0.43	Homogeneous
Oxford-Q	43	485	0.54	Homogeneous

Table 3. Education experiences network of CAS-QC and Oxford-Q

3.3. Collaboration Pattern

The average path length of CAS-QC is 2.68, which indicates that the institute tends to be cohesive. As for Oxford-Q, its average path length is 5.08, which demonstrates that the center belongs to flexible collaboration pattern. It means that the collaboration among members maintains weak correlations in Oxford-Q, there may exist barriers in knowledge flow.

Institutes	Average Path Length	Types
CAS-QC	2.68	cohesive
Oxford-Q	5.08	flexible

Table 4. Collaboration pattern of CAS-QC and Oxford-Q

4. Conclusion

The goal of this study is to explore potential problem in talent group composition of institutes, thus a three-dimensional analysis model is proposed to evaluate talent group composition of institute. According to the indicator of each dimension, age distribution, education experiences, collaboration pattern, research institutes are divided into different types. In empirical study, the potential problems and the differences between the two research centers are revealed by this model.

All the data in this model comes from objective information, and this means whoever wants to use this method may face the data collection challenge. At the same time, the model needs to be adjusted and interpreted according to the actual situations to get reasonable research findings.

References

[1] Williams, K. Y., Iii, C. A. O. (1998). Demography and Diversity in Organizations: A Review of 40 Years of Research, *Research in Organizational Behavior*, 20 (3) 77-140.

[2] Jehn, K. A., Chadwick, C., Thatcher, S. M. B. (1997). To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on workteam outcomes, *International Journal of Conflict Management*, 8 (4) 287-305.

[3] Jun, Shao. (2007). Construct knowledge sharing strategy with balanced team role. *Nanjing University of Aeronautics and Astronautics*.

[4] Lei, Zhao. (2013). Research on the influence of R & D team member role matching on team performance: taking team interaction process as a mediator. *Shandong University of Finance and Economics*.

[5] Linlin, Jin., Dehui, Li. (2012). Research on the effect of scientific research team members' interdisciplinary characteristics on innovation performance: the mediating roles of knowledge sharing and integration, *Studies in Science of Science*, 1, 111-123.

[6] Zeyuan, Liu. (1982). Talking about Talent Group Composition, *Medicine and Philosophy*, (3), 38-39.

[7] Stvilia, Besiki., Charles, C., Hinnant, Schindler, Katy., Worrall, Adam., Burnett, Gary., Burnett, Kathleen., Michelle, M., Kazmer, Paul, F., Marty. (2014). Composition of scientific teams and publication productivity at a national science lab, *Journal of the Association for Information Science & Technology*, 62 (2) 270-283.

[8] Pelled, L. H. (1996). Demographic Diversity, Conflict, and Work Group Outcomes: *An Intervening Process Theory, Organization Science*, 7 (6) 615-631.

[9] Paul, B., Paulus, Timothy, S., Larey, Anita, H., Ortega. (1995). Performance and Perceptions of Brainstormers in an Organizational Setting, *Basic & Applied Social Psychology*, 17 (1) 249-265.

[10] Jehn, K. A., Neale, M. A. (1999). Why Differences Make a Difference: A Field Study of Diversity, Conflict and Performance in Workgroups, *Administrative Science Quarterly*, 44 (44) 741-763.

[11] Men Weili, Zhang Zhiqiang. (2013). Research Review of Age of Peak Creative Performance, *Studies in Science of Science*, 31(11).

[12] Macleod, Roy. (1978). Scientific Elite: Nobel Laureates in the United States, *Physics Bulletin*, 29(10), 476-476.

[13] Lee, D. H., Seo, I. W., Choe, H. C. (2012). Collaboration Network Patterns And Research Performance: The Case Of Korean Public Research Institutions, *Scientometrics*, 91(3), 925-942.

[14] Hansen, M. T. (1999). The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits, *Administrative Science Quarterly*, 44 (1) 82-111.

[15] Cui Aixiang. (2010). Analysis of Complex Network Evolution Model. *University of Electronic Science and Technology of China*.

[16] Yue Zenghui. (2014). Modeling Study on Knowledge Diffusion in Scientific Collaboration Network based on Co-authorship. *University of Chinese Academy of Sciences*.

[17] Backstrom, L., Boldi, P., Rosa, M. (2012). [ACM Press the 3*rd* Annual ACM Web Science Conference - Evanston, Illinois (2012.06.22-2012.06.24)]. *In*: Proceedings of the 3*rd* Annual ACM Web Science Conference on - WebSci \"12 - Four degrees of separation, 33-42.