The study on the core personality trait words of Chinese medical university students based on social network analysis
Ying Wu, PhD*, Yunzhen Xue, PhD, Zhanling Xue, MD

Abstract
The medical university students in China whose school work is relatively heavy and educational system is long are a special professional group. Many students have psychological problems more or less. So, to understand their personality characteristics will provide a scientific basis for the intervention of psychological health.

We selected top 30 personality trait words according to the order of frequency. Additionally, some methods such as social network analysis (SNA) and visualization technology of mapping knowledge domain were used in this study.

Among these core personality trait words Family conscious had the 3 highest centralities and possessed the largest core status and influence. From the analysis of core-peripheral structure, we can see polarized core-peripheral structure was quite obvious. From the analysis of K-plex, there were in total 588 “K-2” K-plexs. From the analysis of Principal Components, we selected the 11 principal components.

This study of personality not only can prevent disease, but also provide a scientific basis for students’ psychological healthy education. In addition, we have adopted SNA to pay more attention to the relationship between personality trait words and the connection among personality dimensions. This study may provide the new ideas and methods for the research of personality structure.

Abbreviations: 16PF = Cattell sixteen personality factor questionnaire, CBF-PI = Chinese big five personality inventory, CPI = California psychological questionnaire, NEO-PI-R = five-factor personality questionnaire, QZPS = Qingnian Zhongguo Personality Scale frequency of occurrence, SNA = social network analysis.

Keywords: Chinese medical university students, personality trait words, SNA

1. Introduction
The medical university students in China whose school work is relatively heavy and educational system is long are a special professional group. Many students have psychological problems more or less.[1,2] So, to understand their personality characteristics will provide a scientific basis for the intervention of psychological health.

The study on personality is important part in psychology field and the personality words that are basic research tools have been widely used in psychology field. The study from the perspective of the vocabulary dimensions already had a long history in the west.

16PF = Cattell sixteen personality factor questionnaire, CPI = California psychological questionnaire, NEO-PI-R = five-factor personality questionnaire, QZPS = Qingnian Zhongguo Personality Scale frequency of occurrence, SNA = social network analysis.

The study on personality is an important part in psychology field and the personality words that are basic research tools have been widely used in psychology field. The study from the perspective of the vocabulary dimensions already had a long history in the west.

2. Methods
In this study, there were in total 118 personality trait words that came from 5 mature psychological scales: NEO-PI-R, 16PF, CPI, CBF-PI, and QZPS.

We tested 261 medical university new students and everyone should choose 10 words reflecting their personality from the 118 personality trait words. We have taken back 256 scales and rejected 5 scales and the effective rate was 98.1%.

Then, we selected 30 words of high frequency after dealing with the date which was imported into the software excel to form co-occurrence matrix (Table 1). Ucinet, Netdraw, and SPSS were used in SNA.

Ethical approval was not necessary for this study, because all the data were obtained from Chinese university freshman’s psychological census data.
3. Results

3.1. Analysis of centrality

Figure 1 was the structure map of 30 core personality trait words. The value was bigger between 2 nodes, the occurrence relationship was stronger and the distance is closer. Otherwise, the value was smaller between 2 nodes, the occurrence relationship was weaker and the distance was farther. The size of nodes represented the level of degree centrality. The thickness of the line represented co-occurrence times of each pair of words. The thicker the line, meant the more co-occurrence times.

We ever applied centrality analysis which reflects status and rights of activities in their social network on scientific collaboration in Chinese psychiatry research. There are 3 common centrality measures: degree centrality, betweenness centrality, and closeness centrality. In this network, degree centrality was equal to the number of nodes that connect with a central node. That was, if a word had the highest degree centrality, it was considered a central node in the collaboration network. Betweenness centrality is the number of the shortest paths that pass through a given node. In our study, the highest betweenness centrality would indicate that a word possesses and controls a great deal of research resource. Finally, closeness centrality of a node is equal to reciprocal of the total distance from this node to all other nodes. It meant the closer a node was to all other nodes, the higher is its closeness centrality. The lowest closeness centrality indicated a word was at the core position of the entire network.

In this network, we selected the top 10 words with the highest centrality in this study (Table 2). These 10 words that were strong in terms of power, prestige, and influence were the core position in this network. Among these words, Family conscious which had

---

Table 1
The top 30 personality trait words according to the order of frequency.

| Rank | Trait word         | Frequency | Rank | Trait word         | Frequency |
|------|--------------------|-----------|------|--------------------|-----------|
| 1    | Family conscious   | 95        | 16   | Sentimental        | 40        |
| 2    | Kind               | 73        | 17   | Patient            | 39        |
| 3    | Friendly           | 71        | 18   | Emotional          | 39        |
| 4    | Independent        | 62        | 19   | Intelligent        | 38        |
| 5    | Responsible        | 56        | 20   | Introverted        | 37        |
| 6    | Human relations valued | 54   | 21   | Confident          | 37        |
| 7    | Imaginative        | 53        | 22   | Principled         | 37        |
| 8    | Optimistic         | 53        | 23   | Outspoken          | 37        |
| 9    | Tolerant           | 49        | 24   | Frugal             | 35        |
| 10   | Sensitive          | 48        | 25   | Grateful           | 35        |
| 11   | Sincere            | 46        | 26   | Disciplined        | 34        |
| 12   | consciously         | 46        | 27   | Strong-minded      | 31        |
| 13   | Helpful            | 45        | 28   | Serious            | 31        |
| 14   | Harmonious         | 44        | 29   | Dependent          | 30        |
| 15   | Communicative      | 42        | 30   | Gentle             | 30        |

---

Figure 1. The structure map of 30 core personality trait words.
the 3 highest centralities possessed the largest core status and influence.

3.2. Analysis of core-peripheral structure

The analysis of core-peripheral structure which distinguishes a series of actors with higher densities (core) and lower densities (peripheral) is the quantitative research on all the networks. The core actors are in a dominant within the exchanging relationship of peripheral actors.[11] We applied this method into the social network analysis of the core personality trait words and found that correlative value was 0.712 (Ucinet 6.0). From Fig. 2, we can see polarized core-peripheral structure was quite obvious. The 30 personality trait words were divided into core words and peripheral words. The core words were Family conscious, Kind, Friendly, Independent, Responsible, Human relations valued, Imaginative, Optimistic, Tolerant, Sensitive, Consciously, and Helpful. The peripheral words were Sincere, Helpful, Harmonious, Communicative, Sentimental, Patient, Emotional, Intelligent, Introverted, Confident, Principled, Outspoken, Frugal, Grateful, Disciplined, Strong-minded serious, Dependent, and Gentle.

3.3. Analysis of K-plex

A K-plex is a maximal subnetwork in which each node is at least connected with other nodes except these K nodes directly within the subnetwork.[12] In the first, determine the condensation degree of subgroup. To exclude the phenomenon of the fewer number of collaboration, “k” was determined to be “2.” It indicated the members connected with others n/C0 times and these subgroups were higher cohesive in which the members were keeping a relative close relationship. If there were 5 members in the first K-plex, there were 3 numbers connecting with others. In this study, if the value of k was 2, there were in total 588 2-plexs, among which there were factors that are Family conscious, Kind, Friendly, Responsible, Human relations valued, and Imaginative in 2-plexs by using Ucinet (Fig. 3).

| Degree                          | Score | Betweenness | Score | Closeness | Score |
|--------------------------------|-------|-------------|-------|-----------|-------|
| Family conscious               | 553   | Family conscious | 2.613 | Family conscious | 29    |
| Kind                           | 402   | Kind        | 2.613 | Kind     | 29    |
| Friendly                       | 371   | Friendly    | 2.613 | Friendly | 29    |
| Sincere                        | 318   | Imaginative | 2.613 | Imaginative | 29    |
| Independent                    | 309   | Responsible | 2.613 | Responsible | 29    |
| Responsible                    | 285   | Human relations valued | 2.613 | Human relations valued | 29    |
| Human relations valued         | 276   | Sincere     | 2.259 | Independent | 30    |
| Tolerant                       | 261   | Sensitive   | 2.110 | Sincere | 30    |
| Imaginative                    | 245   | Tolerant    | 2.007 | Sensitive | 30    |
| Sensitive                      | 244   | Independent | 1.914 | Tolerant | 31    |
3.4. Analysis of principal components

Principal components analysis (PCA) is a multidimensional factor included in the same system for qualitative and quantitative research which is relatively perfect multivariate statistical analysis methods. PCA takes out several comprehensive variables according to the actual need to reflect the original information and simplify the data. In this study, we introduce the co-occurrence matrix of 30 personality traits words into SPSS 19.0, and then run the principal component analysis, principal components sorting according to the personality traits in descending order of frequency. Eigenvalues, Variance, and Cumulative were respectively calculated (Table 3). Based on the Eigenvalues greater than 1, we selected the 11 principal components and the cumulative was 75.08% (Fig. 4). The 11 principal components were Family conscious, Kind, Friendly, Independent, Responsibility, Human relations valued, Imaginative, Optimistic, Tolerant, Sensitive, and Sincere.

4. Discussion

This study applied SNA into the personality trait words from the analysis of centrality, core-peripheral structure, K-plex, and Principal Components. In these core personality trait words, Family conscious, Kind, Friendly, Responsibility, Human relations valued, and Imaginative were in the core position in 4 analyses above. Three words were Independent, Tolerant, and Sensitive in the core position of the analysis on centrality. Only Tolerant and Sensitive were in the core position of the analysis on centrality. Only Sincere was in the core position of the analysis on core-peripheral structure. Independent, Optimistic, Tolerant, Sensitive, and Sincere were in the core position of the PCA.

From the analysis of these core personality trait words, we could learn that medical university students paid attention to human communication and embodied cultural characteristics relying on the relationship first. It probably related with Chinese harmony and the way of relationship. Second, moral sense as a quality of Chinese self-improvement has been paid great attention by society and medical university students also paid attention to

| Component | Eigenvalues | % of variance | Cumulative % |
|-----------|-------------|---------------|--------------|
| 1         | 4.918       | 16.395        | 16.395       |
| 2         | 3.110       | 10.366        | 26.761       |
| 3         | 2.575       | 8.583         | 35.344       |
| 4         | 2.413       | 8.043         | 43.387       |
| 5         | 1.803       | 6.008         | 49.395       |
| 6         | 1.633       | 5.443         | 54.838       |
| 7         | 1.440       | 4.800         | 59.639       |
| 8         | 1.309       | 4.364         | 64.003       |
| 9         | 1.176       | 3.920         | 72.922       |
| 10        | 1.096       | 3.655         | 76.577       |
| 11        | 1.051       | 3.504         | 75.081       |
| 12        | 0.914       | 3.046         | 78.127       |
| 13        | 0.821       | 2.736         | 80.863       |
| 14        | 0.779       | 2.596         | 83.459       |
| 15        | 0.705       | 2.350         | 85.809       |
| 16        | 0.653       | 2.110         | 87.919       |
| 17        | 0.588       | 1.961         | 89.860       |
| 18        | 0.544       | 1.814         | 91.694       |
| 19        | 0.442       | 1.472         | 93.166       |
| 20        | 0.401       | 1.335         | 94.502       |
| 21        | 0.371       | 1.237         | 95.738       |
| 22        | 0.287       | 0.955         | 96.693       |
| 23        | 0.232       | 0.775         | 97.468       |
| 24        | 0.215       | 0.717         | 98.185       |
| 25        | 0.152       | 0.506         | 98.691       |
| 26        | 0.140       | 0.466         | 99.157       |
| 27        | 0.121       | 0.402         | 99.560       |
| 28        | 0.084       | 0.282         | 99.841       |
| 29        | 0.048       | 0.159         | 100.000      |
| 30        | 0.000       | 0.000         | 100.000      |
enhance their self-cultivation. Third, the responsibility as an important work personality trait also brought to the attention of the medical university students, but the existence of negative words suggested the responsibility levels retrained to be further improved. At last, because medical university students have not entered society, they were unavoidably idealized for their own development and sometimes their emotion was unstable. It conformed to the young personality characteristics. This study not only contained the commonness of human personality, but also reflected the unique personality factor of Chinese medical university students.

Personality was all essence of individual behavior and bad personality traits were the risk factors of the psychological disease. So, the study of personality not only can prevent disease, but also provide a scientific basis for students’ psychological healthy education.

In addition, the previous studies on the personality structure mostly focused on the research of the personality dimensions and personality functional assessment questionnaire. In this study, we have adopted SNA to pay more attention to the relationship between personality trait words and the connection between personality dimensions. This study may provide the new ideas and methods for the research of personality structure.

Acknowledgments

The authors thank the participants for their time and willingness to participate in this study. The authors also acknowledge the support for this study from ‘Analysis on Scientific Collaboration and Trends Prediction of Research Fronts in Psychiatry Field (No. 71503152),’ awarded by National Natural Science Foundation of China. This study was also supported by Program for the Outstanding Innovative Teams of Higher Learning Institutions of Shanxi (OIT).

References

[1] Wang WL, Wang Y. The coping style, personality and psychological healthy survey of 276 medical university student. Chin Medical Health J 2003;17:118–9.
[2] Jiang CM, Zhang XP, Jiang LM. The relationship between mental health and stress of medical college students. Chin J Clin Psychol 1999;7:105–6.
[3] Allport GW, Odbert HS. Trait names: a psycho-lexical study. Psychol Monographs 1936;47:41.
[4] Cattel RB, Eber HE, Tatsuoka MM. Handbook for the Sixteen Personality Factor Questionnaire (16PF). IPAT, Champaign, IL:1970.
[5] Tupes E C, Christal R C. Recurrent Personality Factor Based on Trait Ratings. US Air Force, Lackland Air Force Base, TX:1961.
[6] Liu J. An Introduction to School Network Analysis. Social Science Documentation Publishing House, Beijing:2004.
[7] Ma SQ, Jiao C, Zhang MQ. Application of social network analysis in psychology. Adv Psychol Sci 2011;19:755–64.
[8] Wu Y, Jin X. Analysis of scientific collaboration in Chinese psychiatry research. BMC Psychiatry 2016;16:165.
[9] Wasserman S, Faust K. Social Network Analysis: Methods and Applications. Cambridge University Press, London:1994.
[10] Hou HY, Kretschmer H, Liu ZY. The structure of scientific collaboration networks in scientometrics. Scientometrics 2008;75:189–202.
[11] Liu J. Lectures on Whole Network Approach. Shanghai People’s Publishing House, Shanghai:2009.
[12] Seidman SB, Foster BL. A graph-theoretic generalization of the clique concept. J Math Sociol 1978;6:139–54.