International Journal of Database Theory and Application
Vol.9, No.10 (2016), pp.215-228
http://dx.doi.org/10.14257/ijdta.2016.9.10.18

Exploring the Relationship between Creativity and Character based on Online Text Data Analysis*

Hyunju Choi[1], Younchul Choi[2] and Kyungchul Kim[3]

[1]*Institute for early child education, Korea National University of Education*
[1]*hjchoi1119@knue.ac.kr*
[2]*Dept. Early Childhood Education, Korea National University of Education*
Corresponding Author: [2]*eekjunechoi@knue.ac.kr*
[3]*Dept. Early Childhood Education, Korea National University of Education*
[3]*kbrian@knue.ac.kr*

Abstract

The goal of his research was to analyze the meaning of creativity and character based on online text data analysis and identify the relationship between the two. To perform research, online text data was collected for analysis. The collected data was cleaned and then used to analyze the frequency of keyword text, network density, and centrality. The research concluded that both creativity and character have important socio-cultural significances on education. Also, the research showed that, while people tend to include character when they consider creativity, they rarely include creativity when they consider character. It is hoped that the results of this research will contribute to developing ways to integrate creativity and character in education.

Keywords*: creativity, character, online text data*

1. Introduction

Creativity refers to an ability to produce something original that is different from existing things. Creativity has been valued in all societies throughout history because it is directly related to the success of not only individuals but also society, nation and even human race. Creative thinking is becoming even more important in today's rapidly changing world as we are entering the age of information.

However, there is another human attribute that is receiving attention along with creativity– character. In particular, Korean society today is suffering increasing social problems associated with corruption, alienation and violence, the seriousness of which has caught the attention of people. As a result, Korea became legislate a law on character education. Korean government is currently making comprehensive efforts to enable character education in all primary school education, covering elementary, middle and high school educations, including necessary teacher training and educational program developments to effect character education. Koreans tend to see character as something that can be acquired through education, not inborn personal traits.

In the past, study on creativity was focused on a person's divergent thinking such fluency, flexibility, originality and sophistication. Accordingly, there tremendous and continuous efforts made to measure creative abilities of individuals [1-2]. However, today, the focus of researches on creativity is shifting from creative thinking ability of individuals to the socio-cultural context from which creativity emerges [3-5]. Cskiszentmihalyi (1996) one of the foremost experts in creativity research, interpreted creativity in socio-cultural context [3]. He emphasized a need for considering character

* This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2015S1A5B8036254)

ISSN: 2005-4270 IJDTA
Copyright ⓒ 2016 SERSC

and environmental traits as constituent part of creativity in addition to cognitive traits. Cropley (2001) also added ethical traits as part of constituent part of creativity [6], while Runco (1993) stressed the important of emotional factors in creativity [7]. Martin (2012) emphasized ethical creativity [8]. Ethical creativity is an ability to create something that is original and valuable under the ethical realm. The ethical realm emerges when all the realms of human life come together in convergence. At the end, creativity is completed through an individual's interactions with the socio-cultural context that surround him, and such conditions are changing as the times are changing today. What is necessary here is then communication with others, which could be termed "character realm."

Character, like creativity, also tends to change in meaning and components according to changes in time. That is, in reality, the definition of character and its requirements demanded by society tend to vary depending on the age. The United States uses "Character education" to cultivate character in students [9]. China achieves the same goal through its global citizenship training program and Japan has the educational philosophy of "cultivating richness in mind" in students [10]. As you can see, each country has a different character education program and the components of character education today are not necessarily same as the past. The definition of character is changing over time. It is therefore crucial and necessary to recognize such changes and explore the ways to equip the students with the right character to prepare for future. This lies in the realm of creativity. Peterson and Seligman (2004) included creativity as part of character [11].

If you consider the various opinions of scholars as well as socio-cultural text, creativity and character seem to be very closely related. Of course, there should be more discussions on whether character should be included in creativity and vice versa. This research instead focuses on studying what similarities and differences are shared by creativity and character. The meaning of creativity and character changes according to changes in socio-cultural context. In doing so, it takes on universal social importance rather than being meaningful to a limited group of people. Therefore, studying the socio-cultural context on how people perceive the importance of creativity and character provides crucial insight to the nature of the problem in addition to the opinions of the experts in the field.

This research used big data analysis in order to explore how creativity and character are being perceived in socio-cultural context. Big data refers to massive data of various types that are being produced in real time [12]. Mathematical techniques are used to big data to make probabilistic inferences. That is, use of big data allows us to analyze today and make predictions based on it. Big data is making great progresses in politics, economic and public health today. As the value of big data is being recognized more and more, it is being used as effective research method in education and other social sciences. Big data analysis reveals various relationships that exist among data. It allows us to formulate a principle mechanism behind various social phenomenon and use the results to understand the present and predict the future. Therefore, the use of big data was a great way of achieving the objective of our research goal.

The goal of this research is to use big data to explore and identify the meaning of creativity and character in socio-cultural context. The questions asked by this research were: 1) What is the meaning of creativity in socio-cultural context? 2) What is the meaning of character in socio-cultural context? 3) What is the meaning of "creative character" in socio-cultural context?

 Copyright ⓒ 2016 SERSC

2. Research Method

2.1. Data Collection

Text data was collected from online and analyzed in order to find out how creativity, character, and creativity-character are being perceived in socio-cultural context. Data collection was done through Textom [13], data collection solution developed in Korea. The first data collection was done by entering the keywords “creativity” and “character” to Textom search box. The second data collection was done by entering “creativity” in Textom search box and inputting “character” in the detailed search conditions as the term that must be included. For collection channel, Naver café, blog and news; Daum café, blog and news; Google news; and facebook were included. The collection period covered 3 years from August 1, 2013 to July 31, 2016.

Videos and images were not included. Also, professional reports such as research materials, trend report, official statistics, reports/letters or government data were not included. The data collected for this research amounted to 6.4 GB. The detailed description of data is shown in <Table 1>.

Table 1. The Detailed Description of Data

Keyword	*Data Number*	*Data Capacity(KB)*
Creativity	15,965	411,966
Character	16,302	6,171,522
creativity+character	5,611	98,322

2.2. Data Collection

2.2.1. Data Collection and Cleaning Tool

In this research, Textom was used to collect text data. Textom is an integrated big data solution developed by The IMC, Inc. in Korea. Textom is a solution platform that can not only handle data collection but provides data cleaning services as well as simple data visualization services. This researched used the cleaned data provided by Textom (frequency of extracted texts, N_GRAM frequency counts) and matrix data for basic analysis. In addition, the research used the results of visualization provided by Textom (word cloud, bar graphs, network graphs, network word tree) as basic materials for research.

Text data

2.2.2. Social Network Analysis

The term social network was first used by John A. Barnes in 1954. According to John Barnes, social network is a relationship of the patterns among various social variables. Therefore, the analysis of social network tries to explain the effect of social structure by analyzing the inter-relationships among individuals. Since social network analysis focuses on inter-relationships, it is useful to identify various relationships among individuals. In this research, online text big data was used to perform social network analysis in order to understand the relationship between creativity and character in socio-cultural context. Ucinet6 program was used to perform social network analysis. Ucinet program is a free network analysis service to conduct network analysis [14]. In this research, Ucinet6 was used to analyze density, degree centrality, closeness centrality and betweenness centrality. Also, network data was validated for significance.

Copyright ⓒ 2016 SERSC

2.3. Data Analysis Procedure

The data collected by Textom program was subject to first cleaning to make it suitable for the purpose of this research. The first cleaning removed data that is not relevant to this research and the figures, numbers and special character that were unnecessary. Then, the data was downloaded in Excel format and unnecessary data was removed using Excel functions. The cleaned data was then uploaded to Textom program to perform second cleaning via noun cleaning. The final data was then used to conduct word frequency analysis to extract top 50 words. The extracted words were then transformed to symmetric matrix data for social network analysis.

Social network analysis was conducted using Ucinet6 program. Ucinet6 program was used to analyze the network concentration, based on which visualization work was done. Netdraw in Ucinet6 program was used to perform data visualization work. Also, analysis on network density, centrality and significance was done to examine the inter-relationships among the extracted 50 words.

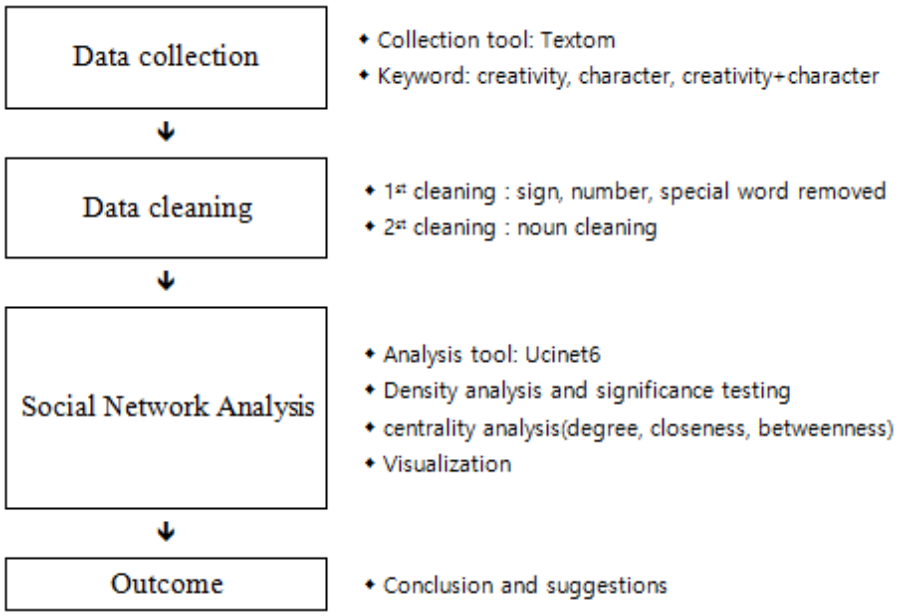

Figure 1. Data Analysis Procedure

3. Conclusions and Discussions

3.1. 『Creativity』 Based on Online Text Data Analysis

3.1.1. Word Number Analysis

On-line text data related to creativity was collected for research. The frequency of keywords in the collected data is shown in <Table 2>.

 Copyright ⓒ 2016 SERSC

Table 2. Keywords Related To 'Creativity'

	Word	Number
1	education	2,787
2	think	1,744
3	person	1,450
4	originality	1,427
5	character	1,390
6	accident	1,297
7	talent	1,273
8	student	1,230
9	element	1,143
10	idea	1,093

According to the frequency of keywords based on co-wording, "education" ranked the highest with 2,787 occurrences, followed by "thinking (1,744)," "person (1,450)" and "originality (1,427)." Also, "character (1,390)" ranked fifth while "talent (1,273)" ranked seventh. In addition, the words such as "ability," "need," "future," and "book" were included as keywords [Figure 2].

Figure 2. Word-Cloud of 'Creativity'

3.1.2. Social Network Analysis

The top 50 keywords analyzed through creativity-related online text data were transformed into a matrix to perform social network analysis. First, the density of network, which indicates the interconnectedness of nodes in a network, was 0.927, suggesting high density of word network in overall data.

Next, significance testing of network data was done. The result shows that the probability of the absolute value of creativity-related network data being greater than Z-score was 0.0002. That is, the relationship among creativity-related network data was statistically significant. Shows the visualization of network analysis of creativity-related words [Figure3].

Copyright ⓒ 2016 SERSC

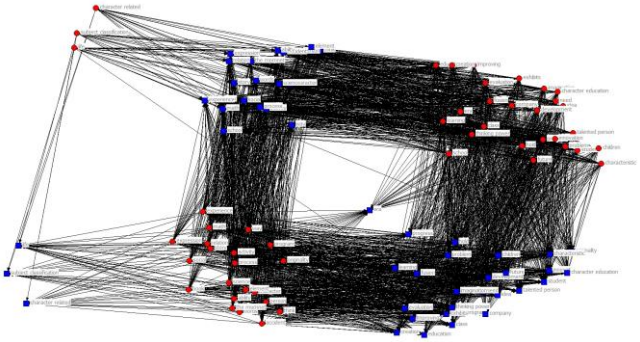

Figure 3. The Visualization of Network Analysis of 'Creativity'

Degree centrality, closeness centrality and betweenness centrality were calculated based on the network of creativity-related words. The result of centrality analysis of the creativity-related words is shown in <Table 3>.

Table 3. Centrality Analysis Based on the Network of 'Creativity'

	Degree Centrality		Closeness Centrality		Betweenness Centrality	
	Word	Centrality	Word	Centrality	Word	Centrality
1	education	6515	experience	100.000	experience	4.627
2	element	3974	think	100.000	think	4.627
3	accident	3949	person	100.000	person	4.627
4	character	3917	story	100.000	story	4.627
5	think	3542	character	100.000	character	4.627
6	originality	3349	accident	100.000	accident	4.627
7	talent	3340	activity	100.000	activity	4.627
8	student	3209	social	100.000	social	4.627
9	person	2895	element	100.000	element	4.627
10	character-book	2607	the moment	100.000	the moment	4.627
11	life	2595	ability	100.000	ability	4.627
12	subject classification	2511	science	100.000	science	4.627
13	need	2357	relation	100.000	relation	4.627
14	science	2350	process	100.000	process	4.627
15	ability	2331	words	100.000	words	4.627
16	program	2257	expression	101.000	expression	4.379
17	idea	2441	originality	102.000	book	4.246
18	problem	1854	way	102.000	math	4.189
19	development	1768	book	102.000	way	1.389
20	way	1680	program	102.000	program	1.389

The result of centrality analysis of creativity-related words show that degree centrality was highest for "education", followed by "element", "accident", "character" and "think". These words are directly related to many keywords related to creativity. In particular, the keywords with high degree centrality had high frequency of occurrence. Therefore, we could infer that the words with high frequency of occurrence tend to have high degree centrality.

Next, the analysis of closeness and betweenness centrality showed that the keywords such as "experience," "think," "person," "story," "character," and "accident" had high closeness and betweenness centrality. High closeness centrality means nodes between keyword network are at a close distance to one another. That is, high closeness centrality implies that the keyword exists at the center of network and also the influence of the

 Copyright ⓒ 2016 SERSC

keyword within the network is high. Also, betweenness centrality allows us to identify the keywords that play the role of mediator between other keywords. Therefore, it can be inferred that the keywords with high closeness and betweenness centrality such as "experience," "think," "person," "story," "character" and "accident" are central to the keyword network related to creativity and play an important role in the network.

What was interesting was that the keywords that had relatively high degree centrality had low closeness and betweenness centrality. In other words, although the keywords such as "education", "talent", "student", "subject classification" were well connected to other keywords within the network, they were outside the center of the network. In contrast, although the keywords such as "experience", "story", "social" had very low degree centrality, they had high closeness and betweenness centrality. It could be inferred that, although these keywords are not well connected to other keywords, they tend to play a central and important role within the network.

3.2. 『Character』 Based on Online Text Data Analysis

3.2.1. Word Number Analysis

On-line text data related to character was collected for research. The frequency of keywords in the collected data is shown in <Table 4>.

Table 4. Keywords Related to 'Character'

	Word	Number
1	character-education	8,320
2	character-story	3,205
3	education	2,142
4	book	1,196
5	student	1,107
6	person	1,073
7	school	878
8	program	854
9	teenager	818
10	character education promotion	814

According to the frequency of keywords based on co-wording, "character-education" ranked the highest with 8,320 occurrences, followed by "character-story (3,205)," "education (2,142)" and "book (1,196)." Also, "student (1,107)" ranked fifth while "school (878)" ranked seventh. In addition, the words such as "object," "progress," "test," and "story" were included as keywords[Figure 4].

Copyright ⓒ 2016 SERSC

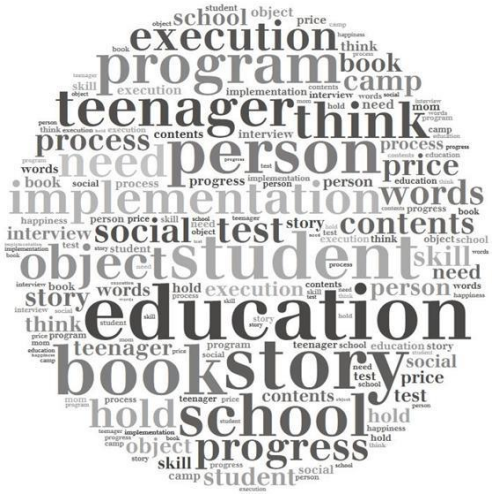

Figure 4. Word-Cloud of 'Character'

3.2.2. Social Network Analysis

The top 50 keywords analyzed through character-related online text data were transformed into a matrix to perform social network analysis. First, the density of network, which indicates the interconnectedness of nodes in a network, was 0.856, suggesting high density of word network in overall data.

Next, significance testing of network data was done. The result shows that the probability of the absolute value of character-related network data being greater than Z-score was 0.0116. That is, the relationship among character-related network data was statistically significant. Shows the visualization of network analysis of character-related words [Figure 5].

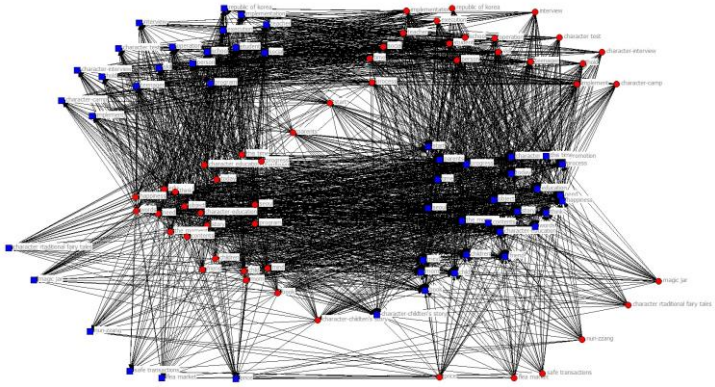

Figure 5. The Visualization of Network Analysis of 'Character'

Degree centrality, closeness centrality and betweenness centrality were calculated based on the network of character-related words. The result of centrality analysis of the character-related words is shown in <Table 5>.

Copyright ⓒ 2016 SERSC

Table 5. Centrality Analysis Based on the Network of 'Character'

	Degree Centrality		Closeness Centrality		Betweenness Centrality	
	Word	Centrality	Word	Word	Centrality	Word
1	character-education	19139	character-education	100.000	character-education	7.701
2	Character children`s story	6346	contents	100.000	contents	7.701
3	education	5785	education	100.000	education	7.701
4	book	3259	need	100.000	need	7.701
5	student	3237	words	100.000	words	7.701
6	school	3135	story	100.000	story	7.701
7	program	2766	think	100.000	think	7.701
8	object	2689	object	100.000	object	7.701
9	teenager	2582	the moment	100.000	the moment	7.701
10	character education promotion	2487	happiness	101.000	happiness	6.743
11	need	2481	children	102.000	children	6.438
12	execution	2406	today	102.000	mom	6.127
13	magic jar	2223	program	103.000	book	5.895
14	character traditional fairy tales story	2114	start	103.000	child	5.868
15	think	2047	mind	103.000	mind	5.868
16	person	2040	progress	103.000	program	5.164
17	implementation	2000	seoul	103.000	seoul	5.164
18	now	1766	parents	103.000	character-story	4.648
19	today	1758	mom	103.000	today	4.612
20	progress	1703	child	103.000	friend	4.089

The result of centrality analysis of character-related words show that degree centrality was highest for "character-education", followed by "character children's story", "education", "book" and "student". Next, the analysis of closeness and betweenness centrality showed that the keywords such as "character education," contents," "education," "need," "words," and "story" had high closeness and betweenness centrality. These keywords are at the center of the keyword network related to "character" and play an important role in it. In particular, the keywords "character education" and "education" had high degree, closeness and betweenness centrality. Therefore, it can be inferred that these keywords are well connected to other keywords in the keyword network, can easily connect to them as well also connect other keywords together. That is, they have such important significance in "character"-related keyword network.

In addition, the result showed that, while "book", "student", and "school" had high degree centrality, they had low closeness and betweenness centrality. Also, while "words," "story," "time" and "happiness" had low degree centrality, they had high closeness and betweenness centrality. Therefore, it is important to examine these words in the network of keywords related to character. Degree centrality shows how wide a word is connected to other words. Closeness centrality shows how closely a word is related to other words. Betweenness centrality indicates how well a word can play the role of connecting one word to another. Therefore, in the network of "character"-related keywords, the keywords such as "words," "story," "time" and "happiness," which can influence other words, are more important than the keywords that are simply connected to other words within the network such as "book", "student", and "school". That is, "words",

Copyright ⓒ 2016 SERSC

"stories", "time", and "happiness" are important keywords that explain the meaning of character in socio-cultural context.

3.3. 『Creativity-Character』 Based on Online Text Data Analysis

3.3.1. Word Number Analysis

On-line text data related to 'creativity+character' was collected for research. The frequency of keywords in the collected data is shown in <Table 6>.

Table 6. Keywords Related to 'Creativity+Character'

	Word	Number
1	Education	1,708
2	character-education	1,160
3	Student	811
4	Talent	791
5	Interview	634
6	Program	417
7	Recharge	403
8	School	402
9	Need	288
10	Think	284

According to the frequency of keywords based on co-wording, "education" ranked the highest with 1,708 occurrences, followed by "character-education (1,160)," "student (811)" and "talent (791)." Also, "recharge (403)" ranked seventh while "need (288)" ranked ninth. In addition, the words such as "future," "evaluation," "activity," and "dream" were included as keywords [Figure 6].

Figure 6. Word-Cloud of 'Creativity+Character'

 Copyright ⓒ 2016 SERSC

3.3.2. Social Network Analysis

The top 50 keywords analyzed through ‘creativity+character’-related online text data were transformed into a matrix to perform social network analysis. First, the density of network, which indicates the interconnectedness of nodes in a network, was 0.958, suggesting high density of word network in overall data.

Next, significance testing of network data was done. The result shows that the probability of the absolute value of ‘creativity+character’-related network data being greater than Z-score was 0.0014. That is, the relationship among ‘creativity+character’-related network data was statistically significant. Shows the visualization of network analysis of ‘creativity+character’-related words [Figure 7].

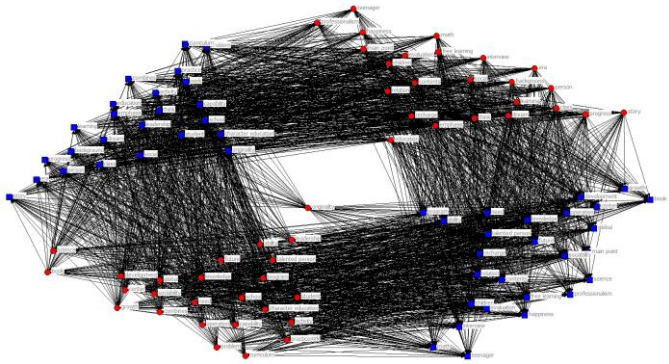

Figure 7. The Visualization of Network Analysis of ‘Creativity+Character’

Degree centrality, closeness centrality and betweenness centrality were calculated based on the network of ‘creativity+character’-related words. The result of centrality analysis of the ‘creativity+character’-related words is shown in <Table 7>.

Table 7. Centrality Analysis Based on the Network of ‘Creativity+Character’

	Degree Centrality		Closeness Centrality		Betweenness Centrality	
	Word	Centrality	Word	Word	Centrality	Word
1	education	4652	education	49.000	education	1.261
2	character-education	3009	character-education	49.000	character-education	1.261
3	talent	2496	student	49.000	student	1.261
4	student	2449	talent	49.000	talent	1.261
5	recharge	1455	leadership	49.000	leadership	1.261
6	school	1402	program	49.000	program	1.261
7	program	1250	recharge	49.000	recharge	1.261
8	future	1031	school	49.000	school	1.261
9	need	1027	need	49.000	need	1.261
10	interview	954	originality	49.000	originality	1.261
11	global	950	ability	49.000	ability	1.261
12	center	902	activity	49.000	activity	1.261
13	operation	860	future	49.000	future	1.261
14	free learning	859	knowledge	49.000	knowledge	1.261
15	dream	846	emphasis	49.000	emphasis	1.261
16	knowledge	829	capability	49.000	capability	1.261
17	ability	799	글로벌global	50.000	social	1.190
18	activity	795	운영operation	50.000	operation	1.169
19	social	790	꿈dream	50.000	practice	1.169
20	capability	754	학습learning	50.000	relation	1.167

Copyright ⓒ 2016 SERSC

The analysis of 'creativity+character' shows that "education," "character-education," "talent", "student", and "recharge" had degree centrality. These keywords were directly connected to many keywords related to 'creativity+character'.

On the other hand, the analysis of closeness and betweenness centrality showed that "education," "character-education," "talent", "leadership", and "program" had high closeness and betweeness centrality. That is, these keywords are at the center of the network of keywords related to 'creativity+character' and play an important role in it. In particular, the keywords such as "education", "character-education", "student", "talent", "recharge", "school" and "program" had high degree, closeness and betweenness centrality. Therefore, it can be inferred that these keywords are well connected to other keywords in the keyword network, can easily connect to them as well also connect other keywords together. That is, they have such important significance in ″character"-related keyword network.

From the overall perspective of examining 'creativity+character' in social context, the related keywords had had similar word frequency and shared similar social network analysis results. The keywords with higher frequency of co-wording ranked higher in each of centrality measures. Therefore, the keywords related to 'creativity+character' at high level all had very important socio-cultural context. However, while "leadership" were low in word frequency and degree centrality, it had high closeness and betweenness centrality, which distinguished them from other keywords. It could be inferred that, although "leadership" did not occur frequently in the 'creativity+character'-related keywords network, it occupied the center of the network and played an important role in connecting other words together.

4. Conclusion

The goal of this research is to survey the meaning of creativity and character in socio-cultural context and identify the relationship between the two using big data. Textom was used to collect online text data and the collected data was transformed into matrix data after data cleaning. The final data was used to preform social network analysis.

The result shows that creativity and character was closely related to the words such as "school," "education" and "student" in socio-cultural context. In addition, the result revealed a high frequency of the co-wording of the words such as "ability," "need," "book", "program" and "practice." That is, there is a high social interests in discovering ways to cultivate creativity and character such as educational program, students, and educational environment. It implies that people look at creativity and character as something that can be acquired through education.

What is interesting about the result of this research is that, the data related to "creativity" included keywords related to "character" and ranked high in all centrality analyses. That is, there are many efforts being made trying to connect creativity and character in socio-cultural context and derive implications. On the other hand, the data related to "character" did not include keywords related to "creativity." That is, creativity was not included as one of the factors that need to be considered when considering character. Not just the term "creativity" was not included, but any terms related to creativity were not included. In fact, while many researchers[3][6][7] who study creativity include character as constituent part of creativity, it is relatively rare for researches who study character to try to explain character in the context of creativity. It appears that the researches in the related disciplines are exerting big influence on socio-cultural consciousness. Therefore, more in-depth researches from the related fields are required on this part. This is because character, also demands creative abilities such as developing social relationships and solving problems under diverse situations. The fact that creativity-related elements are almost completely absent in the character-related data in

 Copyright ⓒ 2016 SERSC

socio-cultural context implies that there is an error in understanding the true meaning of character.

This research identified the differences and the similarities between creativity and character based on online text data analysis and analyzed the relationship between the two, thus providing basic materials for linking creativity and character in future researches. The research results showed that creativity and character are both closely related to education-related elements. It is hoped that the results of this research will facilitate more in-depth researches on creativity and character in the future.

References

[1] H. Eysenck, "Creativity and Personality: An Attempt to bridge divergent traditions", Psychological Inquiry, vol. 4, **(1993)**, pp. 238-246.
[2] E. P. Torrance, "Torrance Test of Creative Thinking: Norms-technical manual", Bensenville, IL: Scholastic Testing Service, **(1974)**.
[3] M. Csikszentmihalyi, "Creativity: Flow and the psychology of discovery and invention", New York: Harper Collins, **(1996)**.
[4] J. C. Kaufman, "Creativity 101", New York: Springer Publishing Company, **(2009)**.
[5] R. J. Sternberg, T. I. Lubart, "Investing in creativity", American Psychologist, vol. 51, no. 7, **(1996)**, pp. 677-688.
[6] A. J. Cropley, "Creativity in education & learning: A guide for teachers and education", London: Kogan Page, **(2001)**.
[7] M. A. Rnuco, "Divergent Thinking, Creativity, and Giftedness", Gifted child Quarterly, vol. 37, no. 1, **(1993)**, pp. 16-22.
[8] M. W. Martin, "Creativity: Ethics and Excellence in Science", MD: Lexington Books, **(2007)**.
[9] T. Akin, "Character Education in American's School", Interchice Publishing Co, **(1995)**.
[10] H. Keung-Ja, "Character Education of Youth", Seoul: Hakjisa, **(2004)**.
[11] C. Peterson, M. Seligman, "Character strengths and virtues: A handbook and classification", London: Oxford University Press, **(2004)**.
[12] L. Douglas, "The importance of 'Big data'", **(2012)**, http://www.gartner.com/technology.
[13] http://textom.co.kr/home/main/main.php
[14] https://sites.google.com/site/ucinetsoftware/home

Copyright ⓒ 2016 SERSC

Copyright ⓒ 2016 SERSC