Image Search based on Words Extracted from Others' Utterances for Effective Idea Generation

Yutaka Yamaguchi[1], Daisuke Shibata[2]
Graduate School of Science and Engineering,
Saga University, Saga, 8408502

Chika Oshima[3], Koichi Nakayama[4]
Faculty of Science and Engineering,
Saga University, Saga, 8408502

Abstract—**People often engage in brainstorming because they want to develop attractive products that involve a new idea. Consequently, many studies, methods, and systems that aim to help people generate ideas have been proposed. We developed the search websites images using search suggestions (SWISS) system, which displays images based on a word extracted from brainstorming participants' utterances and adds additional words using an autosuggest function to stimulate idea generation. We aimed to determine whether the images searched based on the other participants' utterances or those of other participants were more effective for idea generation. Sixteen university students participated in a brainstorming session using SWISS in two conditions. In Condition A, the participants could see the images searched based on the other participants' utterances. These were projected onto a wide display behind each participant during the brainstorming session. In Condition B, the participants could see the images searched based on their utterances, which were displayed on a smartphone. The results indicate that the rate at which the images were related to the ideas in Condition A was higher than in Condition B. SWISS could spread the participants' ideas through the images using an autosuggest function and extract words from the other participants' utterances.**

Keywords—*Autosuggest; brainstorming; search word; smartphone; SWISS*

I. Introduction

The Search Websites' Images using Search Suggestions (SWISS) system [1][2] displays images of search results for participants during a brainstorming session [3] to generate new ideas. Some methods and systems that support idea generation have been developed [4][5][6][7]. Wang [8] stated, "Using pictures as extra stimuli may be more effective than language in stimulating idea generation." Besides, artefacts enhance creativity in innovation tasks [9]. Images, photographs, animations, or video clips of new inventions or strange things enhance viewers' creativities [10]. Thus, Wang [11] developed "IdeaExpander," which is a tool that supports group brainstorming by displaying pictures based on chat conversations. SWISS also displays images based on the utterances of participants in a brainstorming session in real time. Because typical information is almost useless for idea generation [12][13], the images are searched for based on words extracted from the participants' utterances, as well as other words predicted based on a letter of the alphabet or Hiragana via an autosuggest function.

Shibata [1] examined whether images searched for via SWISS allow participants in a brainstorming session to generate an idea more easily than images searched for based only on participants' utterances, without the autosuggest function.

The participants could see the differences between the images being searched for using SWISS and those searched for based only on their utterances. For example, in the former condition, although no participants said the word "car," images of cars were displayed. Then, one of the participants came up with the idea of "a bus for evacuation" as a method that would facilitate elderly people's immediate evacuation from a danger zone.

In contrast, we noted the problem that SWISS could not display images when no participant was speaking, because no words were input to the system. The participants wanted to see the images when they stopped speaking. Therefore, Yamaguchi [2] improved SWISS to display the images searched for based on the last extracted word and re-added the word via the autosuggest function whenever participants stopped speaking. This, however, led to the unresolved problem of the new utterances not being input to SWISS while the images were being displayed. Therefore, in this paper, SWISS is improved to capture new utterances while displaying their related images and to display images when no utterances have been made, based on stored text data of recent utterances.

Moreover, this paper compares the effect of the images searched for using the participant's own utterances for generating new ideas with those of images found using other participants' utterances. In general, when the participant gets an idea from a low-related category, the total number of ideas generated increases [14]. Participants within open networks where information is freely shared will be more creative and have more opportunities to generate new combinations [15]. A wide range of perspectives is more likely to emerge when participants approach idea generation from different angles [16][17]. Therefore, it is believed that the images searched for based on other participants' utterances contribute to generating more ideas than the images searched for using the participant's own utterances alone.

The SWISS system is explained in the next section. In Section III, an experiment is conducted to compare the effect of the images related to the participant's own utterances with those of the other participants' utterances. Then, the results of the experiment are discussed in Section IV. The paper concludes with Section V.

II. System Structure

The SWISS system [1][2] is a smartphone application that captures the utterances of brainstorming participants. Table I shows the application's development environment. Fig. 1 shows the construction of the SWISS system. The participants'

TABLE I. Development Environment

Classification	Specific
OS	Android7.1.1
Terminal	Nexus 5X
Integrated development environment	Android Studio
Development language	Java

TABLE II. Conditions and Themes.

group	1st brainstorming condition	theme	2nd brainstorming condition	theme
1	A	S	B	T
2	A	T	B	S
3	B	S	A	T
4	B	T	A	S

utterances are input into the Android Speech Recognizer and converted to text.

Next, the text data are input to Bing Web Search API. At the same time, text data are preserved in a stored text file. The Bing Web Search API searches websites using this text data. Then, the URL for the website at the top of the search results is extracted.

IBM Watson's Natural Language Understanding (Watson NUL) analyzes the website and extracts four words that represent its features. Watson NLU is a part of the IBM Watson API [18]. NLU extracts metadata (enrichment) from the text data using deep learning. There are seven kinds of extraction: entity extraction, sentiment analysis, category classification, concept tagging, keyword extraction, emotion analysis, and semantic role extraction. The analysis results are returned as JavaScript Object Notation (JSON) text.

Fig. 2 shows the results of the analysis displayed in JSON text. This paper uses "concepts" to extract information from the web site. The "text" shows the words extracted from the web site. Fig. 2 shows only two terms from all extracted words, "computer" and "computer data storage." This paper uses the top four words. "Relevance" refers to the relevance between each extracted word and the web site.

The four words are input to Bing Autosuggest API. Each word is listed with a letter of the alphabet or a Hiragana character, which is selected at random, and receives another word that begins with that letter or Hiragana character from Bing Autosuggest API in return. Fig. 3 shows an example of the autosuggest function. Finally, Google Custom Search searches for images tagged with one of the four words and the word suggested by Bing Autosuggest API.

The autosuggest function adds a word to the extracted word from the website; thus, the added word has some relation to the extracted word. If SWISS adds the word at random, Google Custom Search may not be able to search for images of it, because the two words may be unrelated words that no one has searched for.

Fig. 4 shows the search result images displayed on the screen. One screen displays six images as the results for a pair of words, the extracted word and a word added via the autosuggest function. The screen changes every 20 seconds. The utterances of the participants are acquired even while the images are displayed. After the images for all four pairs of words are displayed, the SWISS system sends the preserved words to storage via Watson NLU. This process is repeated throughout the brainstorming session. Observations of the images are dependent on the participants' intent.

III. Experiment

A. Aim

An experiment was conducted to compare the effect of the images searched for using the participant's own utterances with those of the other participants' utterances. We hypothesized that the images searched for based on the utterances of other participants would contribute to generating more ideas than those stemming from the participant's own utterances. This is because unexpected images may spark the participant's thinking when the others' utterances are different from their own.

B. Method

Sixteen male university students participated in the experiment. They were divided into four groups of four participants each. Each participant wore a small microphone, and the sound was input to a smartphone with SWISS installed. Participants were forbidden to operate either this smartphone or their own smartphone.

Fig. 5 and 6 show the two conditions for the experiment. In Condition A, there were four wide displays mounted behind each participant. These were connected to each smartphone to display their screens' images. Each participant could easily see the wide displays behind the two participants sitting in front of him. In contrast, the participants could not see the smartphones' displays, because they were upside down.

In Condition B, each participant could see the images displayed on each smartphone's screen.

Each group discussed generating a product idea twice under both conditions. The instructions were indicated as follows:

Theme S:
Please generate an idea for a product using zelkova that can be sold at a roadside station considering the following features. Zelkova is a tree with high durability, water resistance, is strong, and elastic. Although its weight is relatively heavy, processing zelkova is easy. However, it is easy for insects to eat. The grain is beautiful and has a sense of quality. Take advantage of these features and consider products that can be sold at roadside stations.

Theme T:
Please generate an idea for a product using porous ceramic that can be sold at a roadside station considering the following features. In a broad sense, "ceramic" is a generic term for inorganic materials that have been heated and baked. Porous ceramics have a light weight, are good heat insulators, and excellent sound absorption / silencing properties. Furthermore, it has the property of repeatedly releasing and adsorbing substances appropriately, due to the differences in the concentration of substances from the outside. The porous quality also
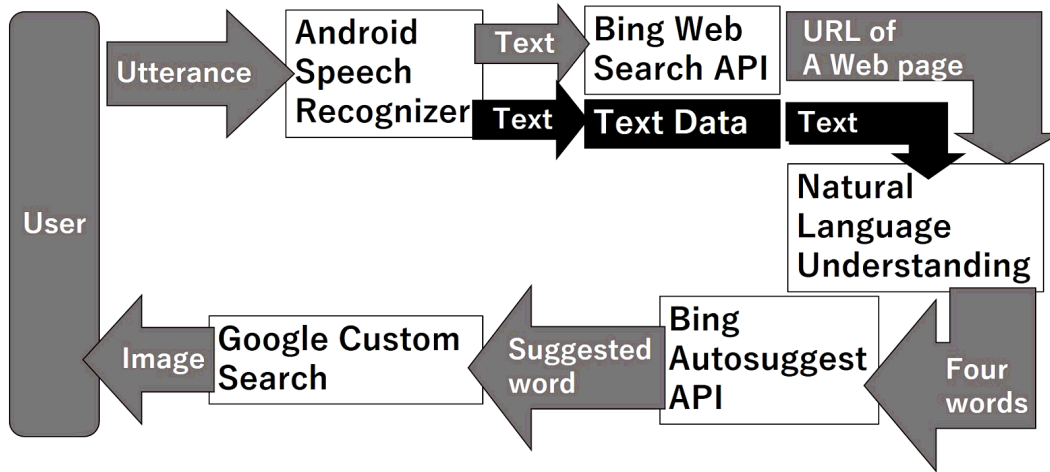
Fig. 1. Flow from inputting utterances to the displaying of images.

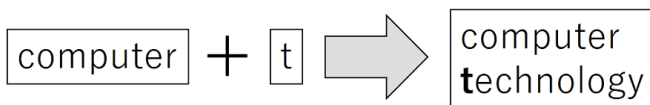Fig. 2. An example of results displayed in JSON text.

Fig. 3. An example of adding a word using the autosuggest function.

Fig. 4. Images of the search results are displayed on the screen.

has a control function regarding the pass-through of specific molecules.

Table II shows the combination of the conditions and the themes in each group.

C. Questionnaire

After two brainstorming sessions, the participants answered three questions in a questionnaire.

Question 1

What did you think of the length of time for displaying the images? (1: very short to 5: very long)

Question 2

How much were you affected by the images when you generated ideas? (1: not at all to 5: very affected)

Question 3

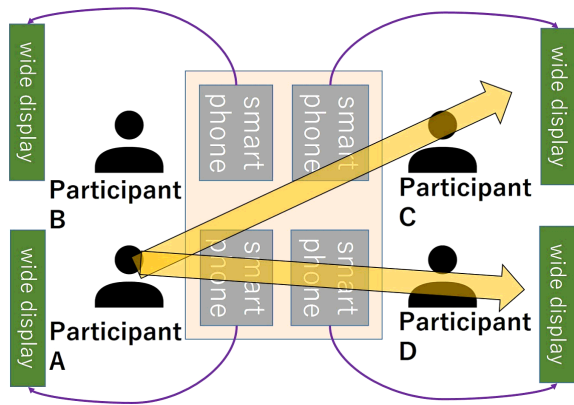Feel free to describe your thoughts about the displayed images.

Fig. 5. In Condition A, the images of the search results are displayed behind each participant.
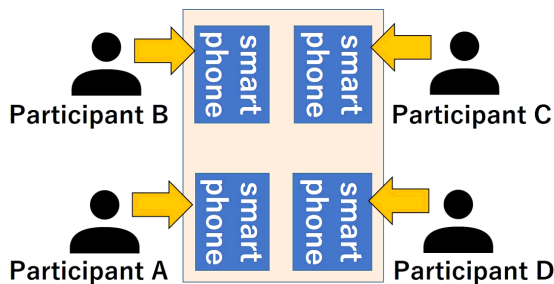

Fig. 6. In Condition B, the images of the search results are displayed on each smartphone.

D. Result

The idea of a "marble," a small glass ball, might be expressed from the images for "chicken egg" (Fig. 7) among the participants of Group 1 during the brainstorming session on Theme S. The following excerpt shows the conversation of the participants before and after coming up with the idea. "[[" means that more than two utterances and/or displays/disappearances of the images occurred. "A–D" means four participants.

Excerpt 1 (English translation):
01B: [[When it comes to wooden toys, I can only think of building blocks.
[[*The images of the search result "chicken egg" were displayed behind Participant D.*
02A: Only building blocks come to mind.
(8 sec.)
03C: Something that looks like a mini-car is made of wood and has wheels.
04A: Oh, I know.
05D: Oh.
(4 sec.)
06B: Something like a toy.
(3 sec.)
07A: [[More beautiful things like marbles, roundly.
[[*Images for the search result "chicken egg" disappeared.*
08B: Process the wood roundly, roundly normally and clean it.

TABLE III. RESULTS OF QUESTION 2.

condition	Ave.	SD
A	2.81	1.05
B	2.56	1.36

No one used the word "chicken" or "egg" during the brainstorming session. However, the search results images for "chicken egg" were displayed on the wide display behind Participant D. Participant B, who was seated in front of Participant D, said, "marbles, roundly." We could not know whether Participant B saw the images before he said this. He did not remember clearly what he had seen during the brainstorming session. People sometimes see images unconsciously. However, the idea of marbles may have been triggered by the images of the round egg.

The average answer for Question 1 in the questionnaire was 3.44. The participants may have thought the length of time while displaying the images was neither long nor short.

Table III shows the results of Question 2. The average was less than 3.0. No significant difference was found among the conditions ($t = 0.54$).

Some answers to Question 3 are shown below.

- Some images were not useful for generating ideas. I also thought that, sometimes, unrelated images were displayed.

- I wish I could set a button to make the images I do not need disappear.

- The images in the first round (Condition A, Theme S) were less related to our conversation than those of the second round (Condition B, Theme T).

- I never looked at the images consciously the first round (Condition A, Theme T). In contrast, it was easy to check the images frequently during the second round (Condition B, theme S).

- In the second round (Condition A, Theme T), the images related to the discussion were displayed, but I could not understand what they were because they were academic contents.

- Especially, in the second round (Condition A, Theme T), the unrelated images tended to appear unless I said keywords that the application could easily extract.

The participants pointed out that some images were not helpful. For example, SWISS displayed images of a convenience store's logo or that of a search engine. Moreover, it was suggested that the conversation among the participants tended to be specialized in theme T. Because the images were academic, it might be difficult to understand what the images were and to generate ideas based on them. It was also determined that the participants needed to look at the images consciously when they were displayed on the wide display.

E. Analysis

1) Aim: It is difficult to establish whether the participants looked at the images on the display or the smartphone. However, if the image was related to the idea that was generated

TABLE IV. THE NUMBER OF IDEAS.

theme	Condition A		Condition B	
	S	T	S	T
group1	26	-	-	15
group2	-	38	40	-
group3	-	19	31	-
group4	18	-	-	21

TABLE V. AVERAGE NUMBER OF COMBINATIONS CONSIDERED RELEVANT.

	theme	
	S	T
Condition A	3.75	8.50
	3.00	3.00
Condition B	4.00	1.50
	3.50	1.25

after displaying the image, it can be assumed that the image contributed to generating the idea. This section examines which images were searched for via keywords based on the participant's own utterances, and which ones were prompted by those of other participants and contributed to generating the idea.

2) Method: First, new ideas from the 16 participants' utterances during the brainstorming sessions were extracted. Table IV shows the number of ideas generated during the experiments. There was no significant difference between the themes ($p = 0.40$) and the conditions ($p = 0.81$), although in Condition A, it was possible to look at the images twice as often as in Condition B.

Second, the images that were displayed two to 30 seconds before starting the utterance were included among the words extracted as being related to the idea. Then, a combination that was consistent with each idea and its applicable images was shown on a sheet, as indicated in Fig. 7.

Four university students evaluated whether even a part of the applicable images was related to each idea. They were instructed in how to make this evaluation according to several examples. If they considered that even a part of the images was/was not related to the idea, the assigned evaluation values were "0" and "1," respectively.

3) Result of Evaluation: Fig. 8 and 9 show the examples of the combinations that the students determined to have a relationship between the images and the idea. In Fig. 8, the images include some dishes. Indeed, it can be said that "tableware" was shown as an idea and displayed among the images. In contrast, in Fig. 9, there was no "figurine" shown clearly among the images.

Table V shows the averages of the number of image combinations that the four students deemed to be related to the idea. A two-way ANOVA was used to compare the mean differences between these factors, the condition and the theme. There was no significant difference between the themes ($p = 1.00$) and the conditions ($p = 0.22$).

Then, rates of ideas related to the images were calculated as follows.

$$r = \frac{RelCom}{NumI}, \quad (1)$$

The idea generated by the participant while brainstorming

marbles, roundly

Images were displayed for 2-30 seconds before starting the utterance that includes "marbles, roundly."

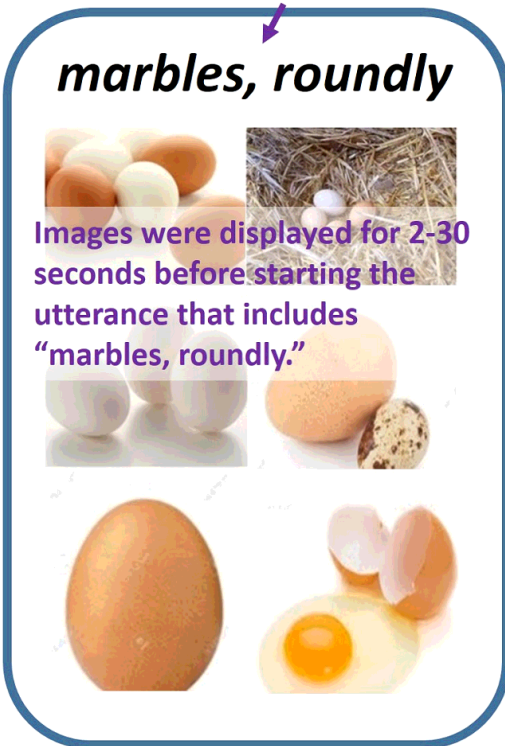These images were searched for in SWISS using the words, "chicken" and "egg."

Fig. 7. A sheet showing the combination of the idea and the images for evaluation.

where r means the rate of ideas related to images, $RelCom$ means the number of combinations considered relevant (see Table V), and $NumI$ means the number of ideas (see Table IV).

Table VI shows the results of the rates of ideas related to images. A two-way ANOVA was used to compare the mean differences between these factors, the condition and the theme. As a result, there was a significant difference in terms of the main effect of the condition ($p = 0.02$). The rate of the ideas that were related to the images in Condition A (17%) was higher than that of Condition B (9%). The result shows the images were searched for via keywords based on other participants' utterances contributed to generating the idea than the images were prompted by those of the participant's own utterances.

IV. DISCUSSION

It is difficult to verify whether each participant came up with his ideas based on images displayed before he spoke. Therefore, an evaluation experiment using four evaluators was conducted to examine the relationship between the idea and the images that the participant who generated the idea might

tableware

Fig. 8. The images were determined to be related to the idea "tableware."

figurine

Fig. 9. The images were determined to be related to the idea "figurine."

TABLE VI. RATES OF IDEAS RELATED TO IMAGES (%).

	theme		
	S	T	Ave.
Condition A	14	22	
	17	16	17
Condition B	10	10	
	11	6	9
Ave.	13	14	13

have seen. The results of the evaluation showed that the rate of images that were related to the ideas in Condition A (the images that were searched for based on other participants' utterances) was higher than in Condition B (the images that were searched for based on the participant's own utterances). The SWISS system displayed images, including unexpected images, by using an autosuggest function [1][2]. Moreover, in this experiment, the participants could see more unexpected images according to searches based on the other participants' utterances. It is known that the other participants' ideas contributed to the generation of ideas [19][20][21]. Similarly, the two features of SWISS in Condition A could contribute to the idea-generation process of the participants.

However, even if an image was determined to be related to the idea, the object, including its idea, was not always included among the displayed images. Therefore, individual differences [22] and cognitive styles [19] are also considerable points to keep in mind, in addition to the participants' possible need
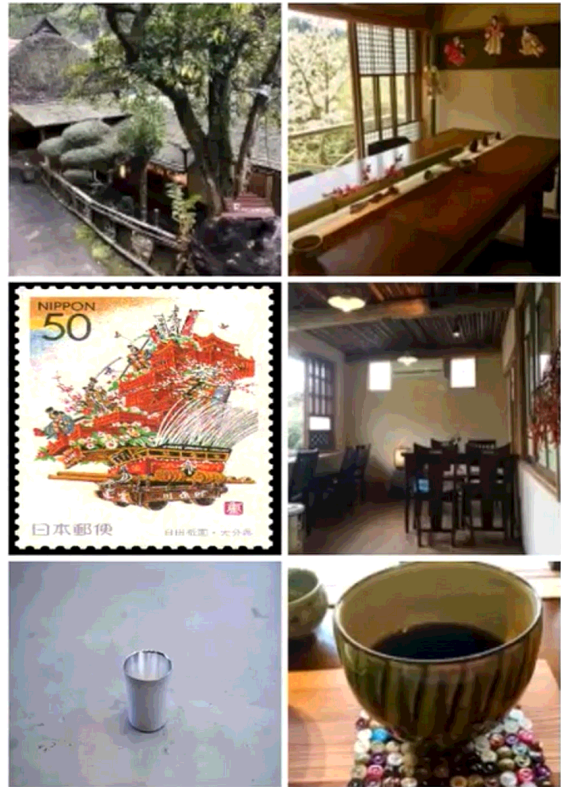
for a fertile imagination [23] to use the images in the most effective way possible.

V. CONCLUSION

SWISS system displays images are searched for based on words extracted from the participants' utterances in a brainstorming, as well as other words predicted based on a letter of the alphabet or Hiragana via an autosuggest function. In this paper, we conducted an experiment to compare the effect of the images that were searched for based on other participants' utterances with the images that were searched for based on the participant's own utterances. The results showed that the former images could contribute to the idea generation process more than the latter images. It was suggested that not only the ideas of others [19][20][21] but also the images that were searched for based on the others' utterances can contribute to the idea-generation process.

In the future, the SWISS system will be improved so that the participants can freely modify the number of images displayed at a time, the size of each image, and the length of time each image is displayed.

REFERENCES

[1]   D. Shibata, Y. Yamaguchi, C. Oshima, and K. Nakayama, Continuous Display of Images Searched by Keywords Extracted from a Brainstorming Session and Suggested by an Autosuggest Function, *Proceedings of the Fourth International Conference on Electronics and Software Science (ICESS2018)*, pp. 76-80, 2018.

[2]   Y. Yamaguchi, D. Shibata, C. Oshima, and K Nakayama, University Students Using SWISS to Display Images Searched by Keywords Extracted from Utterances, *The International Journal of E-Learning and Educational Technologies in the Digital Media (IJEETDM)*, Vol. 4, No. 4, pp. 131-137, 2018.

[3]   A. F. Osborn, Applied imagination: Principles and procedures of creative problem-solving, New York: *Charles Scribner's Sons*, 1953, rev. 1957, 1963.

[4]   H. Ohiwa, K. Kawai, and M. Koyama, Idea processor and the KJ method, *Journal of Information Processing*, Vol. 13, No. 1, pp. 44-48, 1990.

[5]   M. Miura, T. Sugihara, and S. Kunifuji, GKJ: Group KJ method support system utilizing digital pens, *IEICE transaction on information and systems*, Vol. 94, No. 3, pp. 456-464, 2011.

[6]   A. Clayphan, A, Collins, and C. Ackad, B. Kummerfeld, and J. Kay, Firestorm: a brainstorming application for collaborative group work at tabletops, *Proceeding of the ACM international conference on interactive tabletops and surfaces*, pp. 162-171, November 2011.

[7]   S. Tausch, D. Hausen, I. Kosan, A. Raltchev, and H. Hussmann, Groupgarden: supporting brainstorming through a metaphorical group mirror on table or wall, *Proceeding of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, ACM, pp. 541-550, October 2014.

[8]   H. C. Wang, S. R. Fussell, and D. Cosley, From diversity to creativity: Stimulating group brainstorming with cultural differences and conversationally-searched pictures, In *Proceedings of the 2011 ACM conference on computer supported cooperative work*, pp. 265-274, 2011.

[9]   M. J. Eppler, F. Hoffmann, and S. Bresciani, New business models through collaborative idea generation, *International journal of innovation management*, Vol. 15, No. 6, pp. 1323-1341, 2011.

[10]   S. Heinonen, and E. Hiltunen, Creative Foresight Space and the Futures Window: Using visual weak signals to enhance anticipation and innovation, *Futures*, Vol. 44, No. 3, pp. 248-256, 2012.

[11]   H. C. Wang, D. Cosley, and S. R. Fussell, Idea expander: supporting group brainstorming with conversationally triggered visual thinking stimuli, In *Proceedings of the 2010 ACM conference on computer supported cooperative work*, pp. 103-106, 2010.

[12]   T. Montag-Smit, and C. P. Maertz Jr, Searching outside the box in creative problem solving: The role of creative thinking skills and domain knowledge, In *Journal of Business Research*, Vol. 81, pp. 1-10, 2017.

[13]   S. Fu, X. Cheng, T. de Vreede, G. J. de Vreede, I. Seeber, R. Maier, and B. Weber, Exploring idea convergence and conceptual combination in open innovative crowdsourcing from a cognitive load perspective. *n 52nd Hawaii International Conference on System Sciences 2019*, pp. 315-24, 2019.

[14]   V. R. Brown, and P. B. Paulus, Making group brainstorming more effective: Recommendations from an associative memory perspective. *Current Directions in Psychological Science*, SAGE journal, Vol. 11, No. 6, pp. 208-212, 2002.

[15]   L. Fleming, S. Mingo, and D. Chen, Collaborative brokerage, generative creativity, and creative success. *Administrative science quarterly*, SAGE journal, Vol. 52, No. 3, pp. 443-475, 2007.

[16]   F. J. Milliken, C. A. Bartel, and T. R. Kurtzberg, Diversity and creativity in work groups. *Group creativity: Innovation through collaboration*, P. B. Paulus and B. A. Nijstad (eds.) pp. 32-62, 2003.

[17]   B. A. Nijstad, W. Stroebe, and H. F. Lodewijkx, Cognitive stimulation and interference in groups: Exposure effects in an idea generation task, *Journal of experimental social psychology*, Vol. 38, No. 6, pp. 535-544, 2002.

[18]   Watson Natural Language Understanding, https://www.ibm.com/cloud/watson-natural-language-understanding

[19]   M. J. Garfield, N. J. Taylor, A. R. Dennis, and J. W. Satzinger, Modifying paradigms-Individual differences, creativity techniques, and exposure to ideas in group idea generation. *Information Systems Research*, Vol. 12, No. 3, pp. 322-333, 2001.

[20]   B. K. Barnes, Context, framing, and influence, *Industrial and Commercial Training*, Vol. 46, No. 2, 2014.

[21]   H. Hawlina, A. Gillespie, and T. Zittoun, Difficult differences: A socio-cultural analysis of how diversity can enable and inhibit creativity, *Journal of Creative Behavior*, Vol. 53, No. 2, pp. 133-144, 2019.

[22]   Y. Dong, K. M. Bartol, Z. X. Zhang, and C. Li, Enhancing employee creativity via individual skill development and team knowledge sharing: Influences of dual-focused transformational leadership. *Journal of Organizational Behavior*, Vol. 38, No. 3, pp. 439-458, 2017.

[23]   T. Zittoun and F. Cerchia, Imagination as expansion of experience, *Integrative Psychological and Behavioral Science*, Vol. 47, No. 3, pp. 305-324, 2013.