Tan, S. C. et al. (Eds.) (2013). Workshop Proceedings of the 21st International Conference on Computers in Education. Indonesia: Asia-Pacific Society for Computers in Education

Facilitating Creative Cognition by Embodied Conversational Agents

Yugo HAYASHI[a]
[a]*Faculty of Library, Information and Media Science,University of Tsukuba, Japan*
hayashi.yugo.gp@u.tsukuba.ac.jp

Abstract: The study investigated the use of collaborative embodied conversational agents in the facilitation of creative cognition. Based on preliminary studies, two factors were investigated through an experimental design addressing the number of conversational agents (single vs. dual), and method of communication (voice vs. text). 18 participants engaged in a simple interpretation game with embodied conversational agents. Role-playing embodied conversational agents made suggestions on the quality of the participant's interpretations. The study focused on how the two factors enhanced the quality of cognitive process during interactive activities with the agent. Analysis showed that the synergy created by the use of multiple agents along with a voice communication enhanced the cognitive process for the quality of creative interpretations. These results suggest that the number of agents and the method of communication are important factors in designing effective embodied conversational agents in creative activities.

Keywords: Embodied conversational agents, meta suggestions, creative cognition, communication media, social influence

1. Introduction

Past studies in HCI has pointed out the effective usage of embodied conversational agents (ECA) for facilitating learning such as providing meta cognitive suggestions and posing questions (Holmes, 2007; Kim, Baylor & Shen, 2007; Hayashi, 2012 a; Hayashi, 2013 a; Hayashi 2013 b). Recently, there is also a big concern on developing effective collaborative systems during creative activities (Dennis & Williams, 2003). It is a new challenge for ECA researchers to develop such innovative systems for facilitating creative cognitive process. To develop such effective systems, there are needs to implement findings in the field of creative cognition and psychology to design such effective systems (Finke, Ward, & Smith, 1992; Smith, Ward, & Schumacher, 1993). Taken all this, the present study will focus on the use of role-playing ECA's for effective interaction based on models of creative cognition. Especially, the present study will experimentally investigate the two factors (number of ECA's and media type) that are pointed out as effective factors for facilitating interpersonal interaction.

1.1 Creativity model in this study

Studies in cognitive psychology have attempted to understand the cognitive process of creativity. Across the studies, the cognitive model Geneplore has been one of the most reliable means of understanding creative cognition (Finke, Ward, & Smith, 1992). This model posits that creativity is composed of two cognitive processes: (1) a generative phase, in which an individual constructs mental representations, and (2) an exploratory & revise phase, in which those representations are interpreted and used to generate new creative ideas. Figure 1 shows the illustration of the creativity process model in the present study. The interpretative portion of the exploratory phase in the Geneplore model plays an important role in reinforcing ideas, self-reflections, and self-regulation. The quality of a self-generated idea is somewhat dependent on the ability to objectively interpret and evaluate the idea. Studies higher-level cognition such as problem solving show that internal biases, such as a conformation bias occur

as negative constraints (Jhonson-Laird, & Wason, 1977). People tend to search for evidence that supports their own hypotheses or expectations.

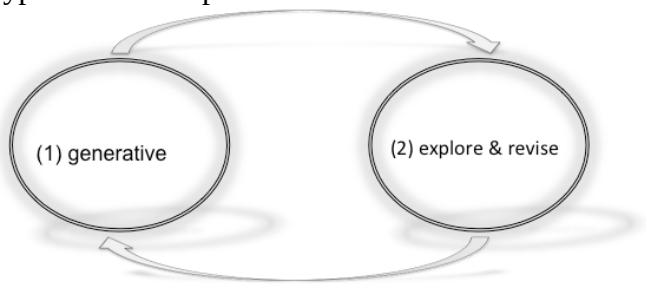

Figure 1. The creativity model in this study.

It is also known that people favor to use knowledge that is accessible to his/her prior knowledge. In creative activities, it is well known that such bias appears and blocks cognitive actives. For example, Smith, Ward, and Schumacher (1993) pointed out the effect of conformity during creative thinking. They used the term fixation to refer to a block or impediment to the successful completion of various types of cognitive operations. For example, fixation can obstruct the retrieval of familiar names or words, such as the names of famous celebrities or politicians. The way that fixation can cause such blocks can also limit the success of creative-idea generation tasks such as divergent thinking and brainstorming. They conducted an experiment about idea generation, in which participants were asked to invent new toys. Prior to the task, one-half of the participants saw three examples of toys that were attributed to fictitious previous participants. Although the three examples were different from each other, each had three critical features in common. The participants who saw the examples used the same critical features, which is indicative of the conformity effect. This study suggests that the when individuals are facing a situation requiring new interpretations, they tend to utilize their available knowledge. To avoid such a bias, it is important to facilitate metacognition, the ability to look at an instance or idea objectively and critically.

Studies of collaborative problem solving in cognitive science have presented results that collaborative partners can play an important role in facilitating the meta-cognitive Okada & Simon, process, which can help avoid cognitive biases (Hayashi, 2012 a). This occurs because feedback from others about the generated interpretations provides an opportunity to rethink the interpretation in a more reflective way (Miyake, 1986; Okada & Simon, 1997). In the present study, it focuses on the type of interaction where a participant generates multiple ideas and is then prompted to make improvements to them. Such interactive activity is an effective strategy in facilitating a objective perspective.

In creative tasks, meta-cognition is important to the formation of better interpretations. However, collaborative activities are costly in terms of manpower, and it is difficult arrange meetings at times that suit all collaborators. To overcome such difficulties, the study focus on the use of ECA as co-partners. The challenge of the study is to investigate the possibility of using these agents in collaborative activities, and discover their potential for increasing the quality of creative cognitive process.

1.1.1 Using multiple ECA's

It is a big challenge for HCI researchers to design an effective ECA that are useful to become peer-collaborators. When considering collaborative activities where conducting creative cognitive activities with agents, it is important to understand the effective presentations towards humans. It is therefore a big concern on understanding that factors on what kind of interactions and designs of agents are useful for influencing human behaviors. Some studies in social psychology has pointed out the effects of 'co-presence' of partners during collaborative activities. It is pointed out that if individuals are experienced in performing a task or expect

they can perform the task well, working in the presence of others impairs performance(Dennis & William, 2003). It is hypothesized that the more the number of the members increase, the more the presence may become stronger and thus facilitate task performance. Then how will the factors such as the increase of the members influence such kind of social facilitation during interaction with an ECA?

Social psychology research has demonstrated the impact of the group dynamics on task performance (Levine, Resnick, & Higgins, 1993). Studies focusing on persuasive communication have shown that the number of other collaborators may influence an individual's decisions. A few studies of human-computer interaction investigated the impact of social pressure from embodied agents. For example, Lee, & Nass (2002). examined the impact of visual representations of multiple agents on performance in a social dilemma task. Beck, Wintermantel & Borg (2005) investigated how social relationships with multiple agents may affect persuasion. Also past studies of the author's studies has investigated the influence of multiplicity during creative cognition (Hayashi, 2012 b; Hayashi, in press). However there are not so many studies that investigate the influence of cognitive process during interactions with such ECA's. These studies imply that under some conditions, agents could motivate and facilitate a change in human opinions. Therefore, agents may play an important role in facilitating differing perspectives through social influence during creative cognitive activities.

1.1.2 Effective communication media

An important point of the interaction with an ECA is the media type, or the method of communication (Joinson, 2001). Adoption of the most suitable method of communication for collaborative interpretation tasks is important to design effective interaction systems. Some studies have focused on the effect of persuasive communication based on the type of media. In these studies, the difference in persuasiveness between text-based communication and oral communication has been the focus (Kiesler, Siegel, & McGuire, 1984). Text-based communication is more likely to change opinions than oral communication. This indicates that the effectiveness of social pressure may differ based on the method of communication.

An interesting survey conducted by Dennis & Williams (2003) shows that activities in electric brainstorming may facilitate creative activities as the number of members increase. It points out that, negative factors known as process-loss can be reduced in text-based interaction compared to oral based interaction. It is also hypothesized that text-based media may enhance the presence of the conversational agents and, thus, create greater social facilitation. The authors conducted experiments based on this line using multiple ECA in a chat-based interpretation task (Hayashi, in press). However, it was still unclear how the quality of the cognitive process may change under such environments. The present study will provide new evidence focusing the cognitive process throughout an interpretation task with an ECA.

1.2 Goal and hypothesis

The goal of this study was to investigate the factors that are useful to the design of embodied conversational agents to be used as peer collaborators. The present study investigates the efficient use of role-taking embodied conversational agents for facilitating creative cognition process during collaborative activities. Specifically, we focus on the influence of two factors that are related to the social psychological and human interface factors: the number of agents and the method of communication. The effects of the number of conversational agents (single vs. dual) and method of communication (voice vs. text) were investigated in an experimental design. The following shows the hypothesis of the present study.

H1: In a creativity task, an individual interacting with several agents will facilitate better cognitive process than a single agent.

H2: Such cognitive process may facilitate when communicating with text-based media compared with oral communication.

2. Method

The present study focuses on the interaction process pointed out in the creative model in Figure 1, and investigates how conformation bias can be eased and unique ideas can be generated with the use of embodied conversational agents. To capture the nature of cognitive process, a nonsensical figure called a droodle was used (Price, 1953). A droodle is a figure that is ambiguous and thus can be interpreted in many ways. Figure 2 shows the ten examples of droodles that were used as the stimuli in this experiment.

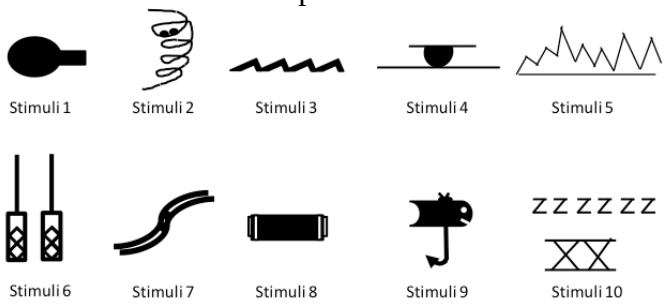

Figure 2. Examples of stimuli used in the study.

The task for the participants in the present study was to generate as many interpretations of the stimuli in limited time. For example, the fifth stimuli (stimuli 5) can be interpreted in many different ways, including mountains, radars, or lightning. Several types of stimuli were taken from the study by Price (1953) and reorganized. In the task, a droodle stimulus was presented to the participants and they were asked to generate as many different interpretations as possible in five minutes. In addition, a conversational agent joined this activity as a collaborating partner, and the participant was required to report his or her interpretations to the agent. The agent's role was to monitor the interpretations and provide suggestions about the generated idea. This part played the role for explore & revise the cognitive process shown in Figure 1. Suggestions were based on an evaluation of the uniqueness of each interpretation. The criteria for the determination of uniqueness was based on a database developed prior to the task (see the next section).

2.1 Droodle database

We conducted a pilot experiment to develop a classification system for the interpretations of the doodles on a unique to non-unique scale. The participants of this experiment, 120 undergraduate university students, were asked to interpret each of the ten stimuli for a free recall test. The experiment took place in a computer room, and students received a course credit. All participants entered their interpretations of the doodles on a web site by inputting their interpretations about the image. The data were collected and pooled for the use as the standard population for general interpretations of the stimuli. For the 10 stimuli in Figure 2, participants generated 244 different types of interpretations, for a total of 1763 interpretations. From each stimuli data set, we calculated the frequency of appearance for each interpretation type, to define its popularity. More specifically, interpretation type $ß$ for stimulus i was coded with the following labels:

- rare: $0\% = ß_i$
- unique: $ß_i \leq 30\%$,
- so-so: $31\% \leq ß_i \leq 50\ \%$,
- major: $51\% \leq ß_i$.

Interpretations that were labeled "major" were considered as easy, "so-so" were considered average, "unique" were considered difficult interpretations to generate, and "rare"

were considered as very creative. From the collected data, 8 out of 10 droodle stimuli, with good category distribution were used in this experiment.

2.2 System

Figure 3 shows the experimental setup. The system was composed of three sub-systems: (1) client interface, (2) server, (3) client agents, and (4) stimuli presenter. The experimental set up followed by the experiment conducted by Hayashi (in press) except with minor changes in parameters of the system.

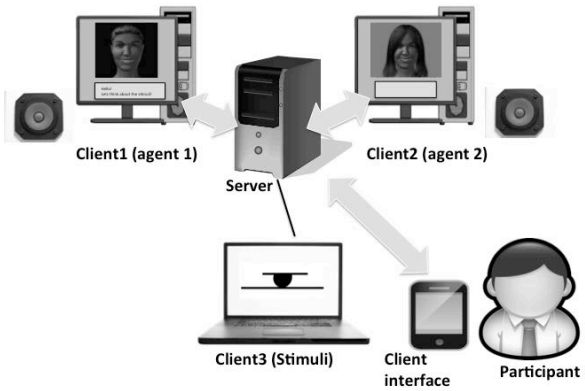

Figure 3. Example of experimental set-up.

The participant was asked to look at the stimuli and make as many interpretations as possible while receiving suggestions from the embodied conversational agents. Participants were told that the agents were collaborating partners and were making suggestions about the uniqueness of their interpretations. Participants used a tablet device (Android OS 4.0) as the communication interface with the computers. A computer server that generated the stimulus of the droodle pictures on a laptop computer was in front of the participant. There were one or two (depending on the experimental condition) client computers with embodied conversational agents presented on an18.5-inch monitor next to the server. All the computers were connected through the local area network (LAN) using TCP/IP. Participants used buttons start the experiment and to change the stimulus presented, and the text field to input their interpretations of the stimulus. Users first clicked the start button and then clicked to view the first picture. Participants were also able to send their interpretations to the other computers by voice or text. Voice recognition API provided by Google was used in the voice option. The server was used for receiving messages from the participant, generating a stimulus, sending information to the agents, and recording the entire process. The stimulus was presented on the server through an application manipulated in Java. The stimulus was presented in the middle of the screen.

The server sent a message to individual clients that signaled them to communicate at the appropriate times. Each client received messages and used a computer-synthesized voice to make suggestions. The client computers had different types of agents installed. The characters that appeared were created using Poser 8 (www.e-frontier.com), which is a design tool for 3D images and animations. The agents' suggestions were generated on the basis of the encoded labels that were stored in the database described in the previous section. When the agent could not detect words from the database, the agent encoded the interpretation as "rare". A typical rule-based system, was used to generate the response by using the rules:

If" major "
-- > "This idea is a popular idea"
ElseIf" so-so "
-- > "Well, this is a little popular response"
ElseIf" unique "
-- > " The idea you say, "%input text%" is quite a nice idea."

Else“ rare ”
-- > “I have never seen such a idea.”

The responses from the agent shown above are examples of several response variations (8 statements) that were randomly selected for each trial. When two agents were present, they alternated making comments to the participant. The timing of sending messages was defined and controlled by the server. The expressions were changed such that they appeared to differ across agents, although meaning remained the same. The agents did not generate any interpretations and only played the role of a collaborative partner by providing evaluative feedback.

2.3 Experimental design and participants

The experiment followed a 2 × 2 between-subjects factorial design (See Table 1). The first factor was the number of the agents that provided suggestions. A single condition was defined as one agent responding, and when two agents responded, it was called a dual condition. The second factor was the method of communication. When text-based messages were used to send comments to the agent, it was called the text condition; when participants used voice messages, it was called voice condition. The participants were instructed to look at the image that was presented on the server, and interpret it. They were instructed to use the tablet device to send messages to the agent. They were told that the agent was a collaborating partner and would respond about the quality interpretation.

Table 1: Experimental conditions.

	single	dual
text	text/voice	text/dual
voice	voice/single	voice/dual

Eighteen undergraduate students participated in the experiment. All participants were assigned to one of the four conditions. Each participant completed all the trials, and the order was counter-balanced. For each condition, there were two trials of stimuli interpretation. Each trial lasted 5 minutes, and the total experiment time was approximately 40 minutes. The experiment took place in an acoustic room, and each participant was alone during the trial. The voice and text conversations of the participants were collected and analyzed. In order to evaluate the creative process, the following variables were calculated by using the evaluation system described in the previous section : (1) 'creative' (rare + unique: $0\% \leq ß_i \leq 30\ \%$), and (2) 'general' (so-so + major: $31\% \leq ß_i$).

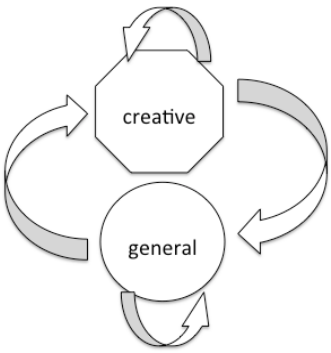

Figure 4. Transitions of the generated words.

The present study focuses on the interaction process of he quality of cognitive process. Therefore, the transitions of the process of the two types of words were analyzed. More specifically, the patterns of the transitions of the following were analyzed during each trial in each condition: (1) general to general, (2) creative to creative, (3) creative to general, and (4) general to creative (See Figure 4.).

3. Results and discussions

Figure 5 lists the average number of interpretations. The vertical axis represents the ratio of the number of generated ideas. Analysis of a 2 × 2 within-subjects factorial ANOVA with the agent number (one vs. two) and communication method (text vs. voice) as independent variables for each transition type. For the type of (1) general to general, there was no significant interaction between the two factors ($F(1, 68) = 1.312$, $p = .26$). However, simple main effects analysis showed that participants in the double-agent condition generated more ideas than did participants in the single-agent condition ($F(1, 68) = 4.593$, $p < .05$). Next, for (2) creative to creative, there was no significant interaction between the two factors ($F(1, 68) = 1.880$, $p = .17$). Also, for (3) creative to general, there was no significant interaction between the two factors ($F(1, 68) = 3.123$, $p = .08$). Finally, for (4) general to creative, there was significant interaction between the two factors ($F(1, 68) = 5.663$, $p < .05$). Next, an analysis of a simple main effect was conducted based on number factor and found that text based interaction was better than oral based interaction in the two agent situation($F(1, 68) = 5.061$, $p < .05$). However there were no differences in other conditions ($F(1, 68) = 1.245$, $p = .27$; $F(1, 68) = 2.165$, $p = .15$; $F(1, 68) = 3.588$, $p = .06$).

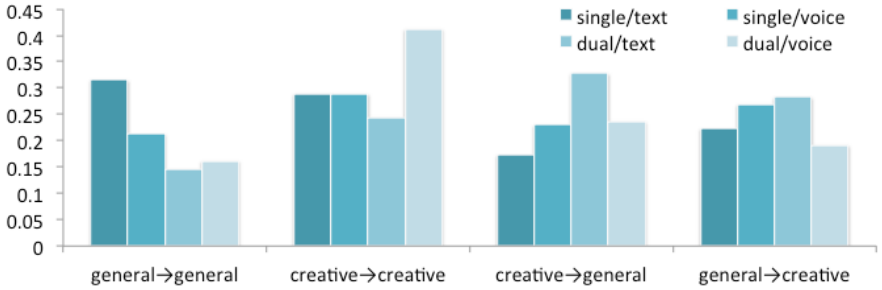

Figure 5. Results of the transitions.

The interaction on the (4) genera-creative show that the synergy created by the use of multiple agents along with a voice communication enhanced the cognitive process for the quality of creative interpretations. These results suggest that the number of agents and the method of communication are important factors in designing effective embodied conversational agents in creative activities. These results both support Hypothesis 1 & 2. This shows that when using ECA's as peer facilitators in a creative generation task, it is effective on facilitating the creative process when using multiple agents along with a text-based interface. These could be important factors on designing creativity supporting collaborating systems in the future.

4. Conclusions

The goal of the study was to investigate the efficient use of role-taking embodied conversational agents in the facilitation of creative cognition during collaborative activities. The study focused on the influence of two factors, the number of agents and method of communication, were related to the social psychological and human interface factors of the peer-partner agents and their human counterparts. These factors were investigated through an experimental design that varied the number of conversational agents (single vs. dual) and the method of communication (voice vs. text). We used a task where participants viewed a stimulus and interpreted it in as many different ways as possible while receiving suggestions from embodied conversational agents. Participants were told that the agents were collaborating partners, and were making suggestions about the quality of the interpretations. Results showed that the use of multiple agents with text-based interfaces facilitates the quality of creative interpretation process. These results suggest that the number of agents and the method of

interaction media is an important factor when designing ECA's as facilitators for enhancing idea generation and reinforcement process in creative cognition.

Acknowledgements

This work was supported, (in part) by 2012 KDDI Foundation Research Grant Program and the Grant-in-Aid for Scientific Research (KAKENHI), The Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXTGrant), Grant No. 25870910

References

Beck, B, U., Wintermantel, M. & Borg, A. (2005). Principles of regulating interaction in teams practicing face-to-face communication versus teams practicing computer-mediated communication," *Small Group Research*, 36, 499–536.

Dennis.A. R., & Williams, M., L.(2003) Electric Brainstorming: Theory, Research, and Future Directions, P.B. Paulus and B. Nijstad, A, eds., no.160-178, *Group Creativity*, Oxford: University Press.

Finke, R. A., Ward, T. B., & Smith, S. M.(1992). *Creative Cognition: Theory, Research, and Applications*. The MIT Press, MA: Cambridge.

Hayashi, Y., & Ogawa, H. (2012 a). Facilitating creative interpretations on collaboration with multiple conversational agents, *Proceedings of the 10th Asia Pacific Conference on Computer Human Interaction(APCHI 2012)*, 443–449.

Hayashi, Y. (2012 b). On pedagogical effects of learner-support agents in collaborative interaction. In S.A. Cerri & B. Clancey (Eds.): *Proceedings of the 11th International Conference on Intelligent Tutoring Systems (ITS 2012)*, Lecture Notes in Computer Science, 7315, 22–32.

Hayashi, Y. (2013 a). Learner-Support Agents for Collaborative Interaction: A Study on Affect and Communication Channels, *Proceedings of the 10th International Conference on Computer Supported Collaborative Learning (CSCL 2013)*, 232–239.

Hayashi, Y. (2013 b). Pedagogical conversational agents for supporting collaborative learning: Effects of communication channels, *Proceedings of the 31th International Conference on Computer Supported Collaborative Learning (CHI 2013)*, 655–660.

Hayashi, Y. & Ono, K. (in press). Embodied Conversational Agents as Peer Collaborators: Effects of Multiplicity and Modality. *Proceedings of the 22nd IEEE International Symposium of Robot and Human Interactive Communication (Ro-Man 2013)*.

Holmes, J. (2007). Designing agents to support learning by explaining. *Computers & Education*, 48(4), 523–547.

Jhonson-Laird, P. N., & Wason, P, C,.(1977) *A theoretical analysis of insight into a resigning task, and post script*. In N. P. Jhonson-Laird and C. P. Wason (Eds.), Thinking: Readings in cognitive science (pp. 143–157). Cambridge: Cambridge University Press.

Joinson, N. A.,(2001). Self-disclosure in computer-mediated communication: the role of self-awareness and visual anonymity, *Journal of Abnormal and Social Psychology*, 31(2), 177-192.

Kiesler, S., Siegel, J,. & McGuire, T. (1984). Social psychological aspects of computer-mediated communication,*" American Psychologist*, 39, 1123–1134.

Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, 23(3), 220–234.

Lee, J, E., and Nass, C.(2002). Experimental tests of normative group inuence and representation effects in computer-mediated communication when interacting via computers differs from interacting with computers," *Human Communication Research*, 28(3), 349-381.

Levine, D., Resnick,B, L. & and Higgins,T. E.(1993). Social foundations of cognition," *Annual Review of Psychology*, 44, 585-612.

Miyake, N. (1986) Constructive interaction and the interactive process of understanding. *Cognitive Science*, 10(2), 151–177.

Okada, T., & Simon, H. (1997). Collaborative discovery in a scientific domain. *Cognitive Science*, 21(2), 109–146.

Price, R. (1953). *Droodles*, Simon & Schuster.

Smith, S. M., Ward,T. B. & Schumacher, J. S. (1993). Constraining effects of examples in a creative generation task," *Memory & Cognition*,21, 837–845.